

AMERICAN JOURNAL of PHYSICS

A Journal Devoted to the Instructional and Cultural Aspects of Physical Science

VOLUME 20, NUMBER 8

NOVEMBER, 1952

Production of Professional Physicists Decreasing

MARSH W. WHITE

The Pennsylvania State College, State College, Pennsylvania

(Received April 22, 1952)

Data are presented to show the numbers of physicists-in-training graduating at various levels during the academic years from 1946-1947 to the present. Following World War II there has been a vast increase in the numbers of physics majors. The number of bachelor's degrees awarded by 525 institutions reached a maximum of 3770 in 1949-1950 and declined each year thereafter, with 3000 estimated for 1951-1952. The number of master's degrees awarded by 158 departments was highest at 1065 in 1949-1950 and has fallen off to an estimated 950 for the current year. At the doctorate level the production is continuing to increase, although at a reduced rate, with 475 estimated for 1951-1952. The total enrollment of undergraduate physics majors reached a maximum of 12,670 in 1949-1950 and has now declined to 10,200. The total number of physics graduate students has been nearly constant for the last two years at 5750.

IN spite of continued shortages and high demand for professional physicists the number being produced at various levels has passed the peak and will continue to decline for some years to come. The enrollment of undergraduate physics majors reached a maximum in the

academic year 1949-1950, and each year thereafter there have been fewer students majoring in physics at this level. The number of graduate students in physics has been relatively stable in recent years, but it is to be expected that they

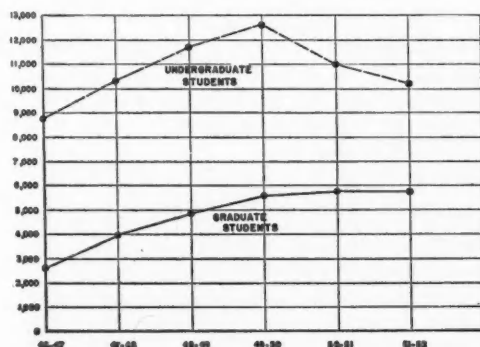


FIG. 1. Enrollment of physics majors after World War II.

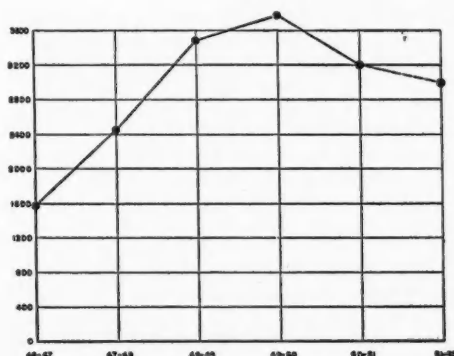


FIG. 2. Bachelor's degrees granted to physics majors after World War II.

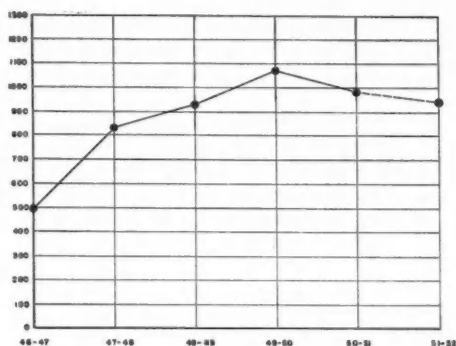


FIG. 3. Master's degrees granted to physics majors after World War II.

also will become smaller because of reduced enrollments and better opportunities for employment (see Fig. 1).

The popularity of physics as a career was greatly accelerated as a result of the spectacular contributions of this science during World War II. Enrollments in the physics major increased by leaps and bounds, with graduates at the bachelor's level reaching a maximum of 3770 in the academic year 1949-1950. For the year 1950-1951 this number declined to 3200 and a still further reduction to 3000 is anticipated for the current year 1951-1952 (see Fig. 2). In spite of this reduced production of physicists at the beginning professional level the current output is vastly higher than the prewar values. The number of institutions offering an undergraduate curriculum with a physics major is now about 550, the largest in the history of the science.

The number of master's degrees awarded in physics in recent years is shown in Fig. 3. Here again is revealed the peak production in 1949-1950 and the substantial reduction of the last

two academic years. The number of institutions offering a program in physics at the master's level has remained much higher than the prewar value and now totals 158.

In examining trends in the production of professional physicists one is naturally mostly concerned with the data showing the development of doctorates. Here the picture remains brighter than in the case of bachelor's and master's degrees. In Fig. 4 it may be seen that the numbers of physics doctorates granted has

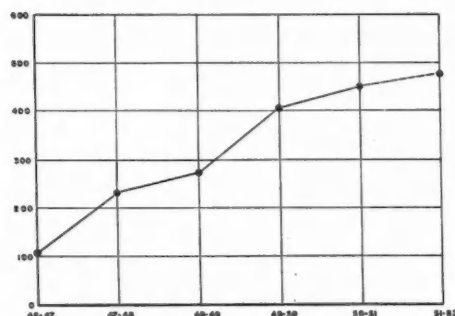


FIG. 4. Doctor's degrees granted to physics majors after World War II.

not yet begun to decline although the reduction in the number of master's and bachelor's degrees granted would indicate that such decrease is imminent. The number of doctorates currently being produced is now sufficient to make substantial inroads upon the deficits caused by the World War II interruption of training of this personnel. In spite of this high productivity, the present and future demand for these graduates is likely to be much larger than the foreseeable supply.

One of the most fundamental laws of terrestrial physics is that of the conservation of mass (energy). The applicability of this law to cosmology forms a very debatable question on which the various theories disagree.

The law itself is clearly an extrapolation from experience. Laboratory experiments show that matter is conserved, at least to experimental accuracy, and for all terrestrial applications this accuracy is amply sufficient. On the other hand, terrestrial regions and periods of time are not typical from a cosmic point of view. The densities are far greater than average cosmic densities, and the periods of time are very short on the cosmic scale. Accordingly, a rate of creation of matter per unit volume per unit time, which is far too low to be detected experimentally, could nevertheless be of the greatest cosmic significance. It follows that terrestrial experience is no guide whatever in assessing whether there are processes of creation or annihilation going on at a rate significant in cosmology, and the question must be left open.—Cosmology by H. BONDI, Cambridge University Press, 1952, p. 73, by permission.

Concerning Statistical Fluctuations in Radioactive Phenomena

ERIC RODGERS AND SARAH H. RODGERS
University of Alabama, University, Alabama

(Received February 18, 1952)

A recent paper by Grundl, Karioris, and Barkow on statistical fluctuations in radioactive phenomena is discussed. It is shown how their theory may be improved. Also, some results of experiments on the same subject by the present authors are presented.

IN an article in the January, 1952 issue of the *American Journal of Physics*, Grundl, Karioris, and Barkow¹ applied the Poisson distribution law to the output of a scale-of-two G-M counter apparatus as recorded on a strip of moving paper. The recording apparatus was connected in series with one discharge tube of the scale of two, and alternate counts from the G-M counter were recorded as tiny dots or pips on the moving paper. They reasoned that the scaler records the first count as zero pips, the second and third as one pip, the fourth and fifth as two pips, etc. They then concluded that if n pips are recorded on the paper in time dt , the number of G-M tube counts could be either $2n$ or $2n+1$. Therefore, they said, the probability of recording n pips in the time dt is the sum of the probabilities of getting $2n$ and $2n+1$ counts in dt . If x is the average number of G-M tube responses that occur in time dt , the probability of m responses in any such time interval dt is, by the Poisson law,

$$W_m = \frac{x^m}{m!} e^{-x}. \quad (1)$$

From this they concluded that the probability W_n of getting n pips in time dt is

$$W_n = \frac{x^{2n}}{(2n)!} e^{-x} + \frac{x^{2n+1}}{(2n+1)!} e^{-x} \\ = \frac{x^{2n} e^{-x}}{(2n)!} \left(1 + \frac{x}{2n+1} \right). \quad (2)$$

If GKB had stopped the counter at the end of each interval dt , and had "zeroed" the recorder before each new interval, their reasoning would have been correct. Since they recorded a large number of pips on continuously moving tape,

and then obtained their intervals by marking off equal lengths on the tape, the Eq. (2) that they used as probability distribution function is incorrect. They overlooked the fact that $2n-1$ is a possibility having for any interval (except the first) the same probability as $2n+1$, and that each of these is just half as probable as $2n$. To see this, consider Fig. 1 which represents an interval AB marked off on the tape. The first pip a in the interval AB can represent 2 counts in AB or 1 depending on which side of the boundary A the first count occurred. Similarly, the last pip b can represent two counts or three counts in AB depending on whether a count occurs between b and the boundary B . The combination 1 with 3 or the combination 2 with 2 both mean $2n$ counts in the interval AB . On the other hand, the combination 2 with 3 and the combination 1 with 2 mean $2n+1$ and $2n-1$ counts, respectively. Therefore, we see that the latter two possibilities are only half as probable as $2n$. Equation (2) used by GKB should, therefore, have been

$$W_n = \frac{x^{2n-1} e^{-x}}{2(2n-1)!} + \frac{x^{2n} e^{-x}}{(2n)!} + \frac{x^{2n+1} e^{-x}}{2(2n+1)!},$$

or

$$W_n = \frac{x^{2n-1}}{(2n)!} e^{-x} \left[n + x + \frac{x^2}{2(2n+1)} \right]. \quad (3)$$

In their experiment, GKB found 3.656 pips per interval as the average for the 1158 intervals

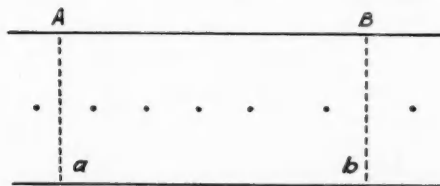


FIG. 1. Sketch of an interval as studied by Grundl and his associates.

¹ Grundl, Karioris, and Barkow, *Am. J. Phys.* 20, 35 (1952). Authors referred to as GKB in text.

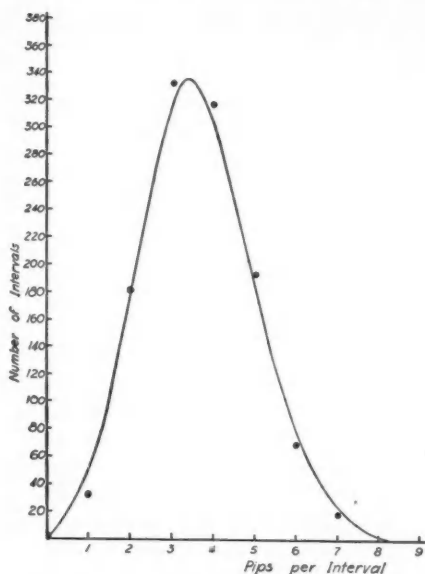
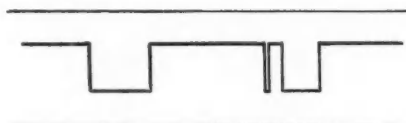
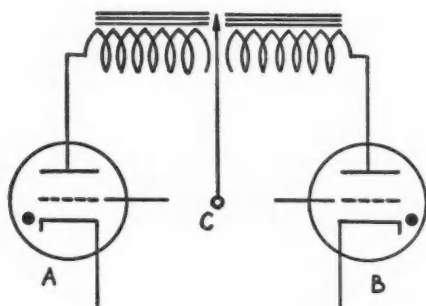


FIG. 2. Our analysis of the data by Grundl and his group. The smooth curve is drawn according to our theory and the indicated points represent their experimental results.



(b)



(a)

FIG. 3. Diagram of our apparatus using a scale-of-two G-M counter circuit. The strip of paper moved just under the pivoted bar *C* and in a direction parallel to the length of *C*. The tracing on the paper appeared as shown enlarged in (b).

measured. The average count \bar{x} per interval is simply double the average number of pips per interval or $\bar{x} = 7.312$. In the present article we have used Eq. (3) in recalculating the probabilities and analyzing the data obtained by Grundl and his co-workers. The results are given in Table I and are shown graphically in Fig. 2. The graph in the article by GKB shows a displacement of the theoretical curve with respect to a curve through the experimental points. This displacement is absent in our Fig. 2.

We have also allowed students to do experiments similar to those described above. In every case, however, we have used apparatus that gave the time distribution of individual counts for which the theory involves only a simple use of Poisson's law. Thus we avoided questions of mutually exclusive events and independent prob-

TABLE I. Comparison of observed data with theories.

Number of pips in an interval	Number of intervals containing \bar{x} pips		
	Experiment	Our theory	GKB theory
0	2	3.6	6
1	33	49	72
2	182	184	228
3	333	317	334
4	318	306	284
5	194	188	155
6	70	79	59
7	19	24	17
8	4	5.7	3

abilities. Several years ago the present authors did some careful experiments very similar to those by Grundl and his group. We used a scale-of-two scaler but had an electromagnet in series with each discharge tube as shown in Fig. 3. As the discharge shifted between tubes *A* and *B*, a pivoted bar of iron *C* was pulled alternately from one coil to the other. A pen attached to the end of *C* made a line on a strip of paper that moved at constant speed. As the discharge shifted in the tubes, the line made by the pen shifted as shown in Fig. 3(b).

In our experiments we recorded 20,000 background counts on a continuous strip of adding machine paper and then carefully marked the total length into 5000 equal intervals. This automatically made the average 4 counts per interval. Our counting rate was low enough to insure that counts lost due to finite resolving time of the

apparatus were negligible. We applied the Poisson formula to obtain the (previously unpublished) results shown in Table II. If we should show these results graphically the curve would look very much like Fig. 2 but with the relative fitting better because of the larger number of intervals.

An alternative method of analysis deals with the distribution of distances on the paper between counts. If k is the average number of counts per centimeter of length on the paper, then ks is the average number for a length s . By the Poisson formula, the probability of a length extending a distance s without a count is

$$W_0 = e^{-ks}.$$

The probability that a length may extend to s without a count and that then at least one count will occur in the following element of length ds is

$$W = e^{-ks}(1 - e^{-kds}). \quad (4)$$

TABLE II. Analysis in terms of counts per interval.

Count per interval	Frequency observed	Frequency predicted	Count per interval	Frequency observed	Frequency predicted
0	98	92	7	296	298
1	352	366	8	172	149
2	777	733	9	74	66
3	962	977	10	25	25
4	972	977	11	14	10
5	764	781	12	3	3
6	491	520	13	1	1

In order to use Eq. (4) to test our data, we measured a total of 15,000 distances between consecutive counts on the paper. We chose for s the values 0.0 cm, 0.5 cm, 1.0 cm, 1.5 cm, etc., in steps of 0.5 cm, and for ds we took the constant value 0.5 cm. The s value resulting from measurement of any interval was recorded as the greatest multiple of 0.5 cm contained in the interval. For example, a length of 1.8 cm between consecutive counts would be recorded as 1.5 cm for s . This would mean that the interval extended 1.5 cm without a count, but that a count came in the following element $ds = 0.5$ cm. Similarly, distances of less than 0.5 cm were recorded as 0 for s , distances between 0.5 cm and 1.0 cm were recorded as 0.5 cm for s , etc.

The results of this analysis are given in Table III and are shown graphically in Fig. 4. The value of k was determined by measurement to be

TABLE III. Analysis of distances between counts.

s (cm)	Frequency observed	Frequency predicted	s (cm)	Frequency observed	Frequency predicted
0.0	3475	3515	5.5	210	187
0.5	2696	2691	6.0	144	143
1.0	2031	2060	6.5	102	109
1.5	1536	1577	7.0	98	84
2.0	1197	1208	7.5	82	64
2.5	910	925	8.0	59	49
3.0	723	709	8.5	43	38
3.5	540	542	9.0	37	29
4.0	420	416	9.5	29	23
4.5	338	318	10.0	19	17
5.0	263	244	10.5	11	13

0.5341 average counts per centimeter of length. This means that the average distance between counts was nearly 2 cm. At first thought the student might expect the graph in Fig. 4 to rise to a maximum in the neighborhood of $s = 2.0$ cm. A closer look at our method of analysis shows why the curve never rises. On the strip of paper we take a count as origin and divide the distance following it into 0.5-cm subintervals. The chance of finding a count in any one subinterval is the same for all subintervals. Suppose we find a count in the first subinterval. We would record the distance as $s = 0$ and would then take this count as a new origin and again mark off the distance into 0.5-cm subintervals. But this procedure denies any other subinterval a chance for any measurement yielding a count in the first subinterval. Therefore, the first subinterval has more chances than any other. Likewise, the second subinterval has more chances than any subinterval beyond it, etc. Consequently, the curve must always drop.

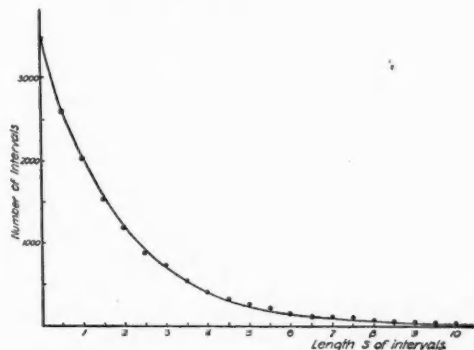


FIG. 4. Graphical comparison of our data with theory. The smooth curve represents Eq. (4) and the indicated points represent experimental values.

Quantum Effects in the Radiation from Accelerated Relativistic Electrons

L. I. SCHIFF

Stanford University, Stanford, California

(Received March 3, 1952)

The radiation associated with the operation of a betatron or synchrotron, due to the transverse acceleration of relativistic electrons by the magnetic field of the machine, is discussed qualitatively. Order of magnitude expressions for the angular opening of the cone within which the radiation is emitted, and for the wavelength of the maximum of the radiated spectrum, are obtained by means of qualitative arguments based on classical physics. It is then shown, from a consideration of the rate of spread of a wave packet that describes the electron, that quantum effects are negligible so long as the ratio γ of the total energy to the rest energy of the electron is small in comparison with the square root of the ratio of the orbit radius R to the reduced Compton wavelength λ_c of the electron. This criterion for the validity of classical theory is substantially equivalent to, but not identical with, the commonly accepted criterion $\gamma^2 \ll R/\lambda_c$, which is derived from the assumption that the momentum of an emitted photon must be small in comparison with the momentum of the electron. Thus classical theory is expected to be valid for electron energies small in comparison with 10^{15} ev if the magnetic field strength does not exceed 10^4 oersted.

I. HISTORICAL INTRODUCTION

THE radiation due to the transverse acceleration of relativistic electrons is responsible for a significant energy loss in large betatrons and synchrotrons. The importance of this effect as a limiting factor in determining the maximum size of a betatron was first pointed out by Iwanenko and Pomeranchuk.¹ In his early consideration of the synchrotron, McMillan² realized that this radiation could be enhanced by partial coherence (constructive interference) between electrons due to the circumferential bunching that is characteristic of synchrotron operation. It was shown by Arzimovich and Pomeranchuk³ and by Schwinger⁴ that most of the radiated energy lies in very high harmonics of the frequency of orbital rotation of the electron,⁴ for which bunching is not expected to result in appreciable coherence. Coherence will greatly increase the energy radiated in the low harmonics, but it was shown by the author⁵ that this loss can be strongly reduced by shielding. Arzimovich and Pomeranchuk³ also pointed out that the radiation is

mainly confined to the forward cone of angular opening mc^2/E radians centered on the instantaneous direction of motion of the electron, where c is the speed of light, and m and E are the rest mass and total energy of the electron. All of these results, for the most part obtained independently of other work, were discussed briefly by the author⁵ and by Schwinger;⁶ more recently, Schwinger⁷ published a detailed treatment of most aspects of the radiation problem.

The work referred to above was based entirely on a classical description of the motion of the electron. The commonly accepted criterion for the validity of classical theory is that the momentum of an emitted photon be small in comparison with the momentum of the electron.⁷ This criterion is easily satisfied by all betatrons or synchrotrons now in existence or contemplated. Quite recently, Parzen⁸ has attempted to show that quantum effects are significant for energies attained by existing machines. However, Judd, Lepore, Ruderman, and Wolff⁹ have found

¹ D. Iwanenko and I. Pomeranchuk, *Phys. Rev.* **65**, 343 (1944).

² E. M. McMillan, *Phys. Rev.* **68**, 144 (1945).

³ L. Arzimovich and I. Pomeranchuk, *J. Phys. (U.S.S.R.)* **9**, 267 (1945).

⁴ This conclusion is also inherent, although it was not explicitly recognized at the time, in the work of G. A. Schott, *Electromagnetic Radiation* (Cambridge University Press, Cambridge, 1912), pp. 109, 110.

⁵ L. I. Schiff, *Rev. Sci. Instr.* **17**, 6 (1946).

⁶ J. Schwinger, *Phys. Rev.* **70**, 798 (1946).

⁷ J. Schwinger, *Phys. Rev.* **75**, 1912 (1949).

⁸ G. Parzen, *Phys. Rev.* **84**, 235 (1951).

⁹ Judd, Lepore, Ruderman, and Wolff, *Phys. Rev.* **86**, 123 (1952). *Note added in proof:* See also H. Olsen and H. Wergeland, *Phys. Rev.* **86**, 123 (1952). More recently, G. Parzen (to be published) has acknowledged some of these errors, but still obtains a limiting energy for the validity of classical theory that is much smaller than that obtained from Eq. (4) below.

mathematical errors in Parzen's work that invalidate his main conclusions.

The principal objective of the present paper is to show by means of a wave packet argument that the validity criterion for use of the classical theory given above is substantially but not exactly correct (Sec. III). As a preliminary, and mainly for pedagogical purposes, order of magnitude expressions for the angular opening of the cone within which the radiation is emitted, and for the wavelength of the maximum of the radiated spectrum, are obtained from simple classical arguments (Sec. II).

II. ANGULAR AND FREQUENCY CHARACTERISTICS OF THE RADIATION

Angular Characteristics¹⁰

The curvature of the path of the electron can be neglected for a qualitative discussion of the angular dependence of the radiation. We assume that the electron is disturbed in some way so that it continually emits radiation. The magnitude of the potential radiated by the moving charge is larger in the forward direction than in other directions because the emitted electromagnetic wave front moves only slightly faster than the electron and is continually reinforced in the forward direction by subsequently emitted wave fronts. This effect may be looked at in the following way (see Fig. 1). Suppose the electron to be an extended region of charge located at e and moving along the line AB with speed v . A distant observer at P measures the potential accumulated by a spherical surface CD that collapses in on P with the speed of light c . The time that is required for this surface to traverse the charge e is a measure of the strength of the potential that later arrives at P . This time is inversely proportional to the difference between the collapse speed of the surface and the component of the electron velocity along the radial line eP , or to $c - v \cos \theta$. This dependence on the angle θ of observation will hold even for the limit of a point charge.

The relativistic relation between energy and speed may be written

$$E/mc^2 = 1/(1 - \beta^2)^{1/2} \equiv \gamma, \quad \beta \equiv v/c.$$

¹⁰ The author is indebted to W. K. H. Panofsky for valuable comments on this section.

Since β is very close to unity and we shall only be interested in small angles θ , we can put approximately:

$$c - v \cos \theta \cong c(1 - \beta + \frac{1}{2}\beta\theta^2) \cong \frac{1}{2}c(\theta_0^2 + \theta^2), \quad \theta_0 \equiv 1/\gamma.$$

Thus the potential that arrives at P at a particular time is measured by the value of $1/(\theta_0^2 + \theta^2)$ at the earlier time when the spherical surface traversed the charge. This means that the radiation is most intense inside a cone with angular opening of order θ_0 that is centered on the instantaneous direction of motion of the electron. For a 300-Mev electron, $\theta_0 \cong 0.1^\circ$, and in general θ_0 is inversely proportional to E .

Frequency Characteristics

The radiation that is emitted by the electron during a time interval t passes by a distant ob-

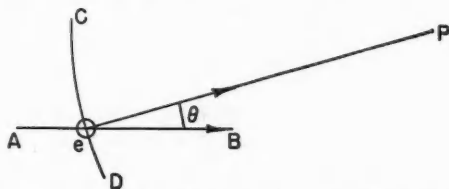


FIG. 1. An extended electron e moves along the line AB with speed v . The potential at a distant observer P is measured by the accumulation of contributions collected by the spherical surface CD as it traverses e while collapsing radially with speed c .

server in the forward direction in a time interval that is much shorter than t because of the relativistic Doppler effect. In Fig. 2, the radiation emitted when the electron is at A lies on the spherical surface CD at the time $t+T$ later, and the radiation emitted when the electron is at B (the distance AB is equal to vt) lies on the spherical surface EF at the time T later. Thus the radiation emitted by the electron as it moves from A to B lies in the shell contained between the surfaces CD and EF at the later time. For a distant observer ($T \gg t$), we can put to good approximation, where BG is perpendicular to AC :

$$\begin{aligned} CE &= AC - AG - GE \cong AC - AG - BE \\ &= c(t+T) - vt \cos \theta - cT \\ &= ct(1 - \beta \cos \theta) \cong \frac{1}{2}ct(\theta_0^2 + \theta^2). \end{aligned} \quad (1)$$

Thus within the forward intensity maximum ($\theta < \theta_0$), the thickness of the pulse that passes

the observer is less than the distance of travel of the electron while the pulse was being emitted, by a factor of order $\theta_0^2 = 1/\gamma^2$.

So far, we have not specified the disturbance that causes the electron to radiate. We now assume that it is a transverse acceleration, which makes the instantaneous path a circular arc of radius R centered at O (see Fig. 3). We consider a segment AB of this arc that subtends the angle $2\theta_0$ at the center. From the preceding discussion, it follows that the central ray AR_2 from the point A is tangent to the arc at A , and the radiation is emitted mainly between the rays AR_1 and AR_3 , each of which makes the angle θ_0 with the central ray. Similarly, the extreme rays BS_1 and BS_3 make angles θ_0 with the central ray BS_2 which is tangent to the arc at B . Then it is apparent from Fig. 3 that AR_3 and BS_1 are parallel (they are actually coincident, although not drawn that way to avoid confusion in reading the drawing). We station an observer a long distance away in this direction. The observer then sees radiation that arises only from the segment AB , since only then is he within the electron's forward cone of angular opening θ_0 .

We have seen that the thickness of the pulse of radiation at a distant point is roughly equal to ct/γ^2 , where t is the time of transit of the electron from A to B . From Fig. 3, this time is found to be roughly equal to $R\theta_0/c = R/c\gamma$, so that the radiation pulse has a thickness of order R/γ^3 when it passes the observer. The duration in time of the pulse at the observer is thus of order $R/c\gamma^3$, and a Fourier analysis of this pulse will give angular frequencies extending from very low values up to about $c\gamma^3/R$. This means that the radiation will consist of reduced wavelengths

(wavelength divided by 2π) extending from very large values down to about

$$\lambda_0 = R/\gamma^3. \quad (2)$$

For a 300-Mev electron in a magnetic field of 10^4 oersted, $R \cong 100$ cm, and $\lambda_0 = 2\pi\lambda_0 \cong 300\text{\AA}$. In general, λ_0 is inversely proportional to E^3 for fixed R , and inversely proportional to E^2 for fixed magnetic field strength.

III. VALIDITY CRITERION FOR THE USE OF CLASSICAL THEORY

Classical theory assumes that the electron is a point charge, and so can only be valid when the quantum-mechanical calculation can be cast in such a form that the electron behaves like a localized charge. Localization to a point is however unnecessary, as well as impossible. It is sufficient that it be possible to construct a quantum-mechanical wave packet such that each of its dimensions is small in comparison with the reciprocal of the corresponding component of the propagation vector of the emitted radiation. In the present case, the main part of the radiation is emitted within the angle θ_0 of the forward direction with propagation vector of magnitude $1/\lambda_0$. Thus the longitudinal dimension of the wave packet must be small in comparison with λ_0 , and the transverse dimensions must be small in comparison with λ_0/θ_0 .

If we consider, for the moment, just the longitudinal dimension, we see that the wave packet cannot be made smaller than the reduced de Broglie wavelength $\hbar/mc\gamma$ of the electron, since otherwise shorter wavelength components would be required, and the momentum and energy of the electron would be increased. Then the condition

$$\hbar/mc\gamma \ll \lambda_0$$

is equivalent to the condition

$$\hbar/\lambda_0 \ll mc\gamma,$$

which is the usual criterion that the momentum of each emitted photon be small in comparison with the momentum of the electron. With the help of Eq. (2), we immediately obtain:

$$\gamma^2 \ll R/\lambda_c, \quad \lambda_c \equiv \hbar/mc \cong 4 \times 10^{-11} \text{ cm}, \quad (3)$$

where λ_c is the reduced Compton wavelength of

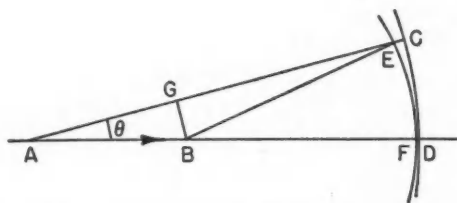


FIG. 2. An electron moves from A to B with speed v in time t . At the time $t+T$ after the electron is at A , or time T after it is at B , the radiation it emits between A and B lies between the spherical surfaces CD and EF . AC is equal to $c(t+T)$, BE is equal to cT , and BG is perpendicular to AC .

the electron. For a magnetic field of 10^4 oersted, this means that E must be small in comparison with 2×10^{16} ev in order for classical theory to be valid. In general, this limiting energy is inversely proportional to the magnetic field strength.

The discussion of the preceding paragraph is defective in that no mention is made of the rate at which the wave packet spreads. The following treatment takes this factor into account, and also considers the transverse dimensions of the packet as well.

We adopt a rectangular coordinate system in which the x axis is along the instantaneous direction of motion of the electron, and the xz plane is the instantaneous plane of motion in which the center of curvature lies. The wave packet must maintain the dimensions given above throughout the time interval over which the radiation process takes place. We have seen in Sec. II that this time is roughly equal to $R\theta_0/c = R/c\gamma$, since only radiation given off during a time interval of this length reaches any one observer. Thus the x dimension of the packet must be small in comparison with $\lambda_0 = R/\gamma^3$ for the time $R/c\gamma$. In addition, the y and z dimensions of the packet must be small in comparison with $\lambda_0/\theta_0 = R/\gamma^2$, also for the time $R/c\gamma$. We now examine each of these three dimensions of the wave packet separately.

Longitudinal Dimension of the Wave Packet

Let δx be the initial longitudinal spread of the wave packet. Then the uncertainty principle requires a longitudinal momentum spread $\delta p_x \sim \hbar/\delta x$. Now $p_x = mc\beta/(1-\beta^2)^{1/2}$, and differentiation shows that $\delta p_x = mc\delta\beta/(1-\beta^2)^{3/2}$. Thus the spread in longitudinal speed is:

$$\delta v_x = \delta p_x / m\gamma^3 \sim \hbar / m\gamma^3 \delta x.$$

This produces an additional spread in the x dimension of the packet which is equal to the product of δv_x and the emission time:

$$\delta v_x R / c\gamma \sim \hbar R / mc\gamma^4 \delta x.$$

Thus if we assume, as was done in deriving Eq. (3), that $\delta x \sim \hbar / mc\gamma$, we find that the packet spreads by the amount R/γ^3 during the emission time. This makes it as large as the radiated wavelength λ_0 , which implies that classical theory cannot be used. The difficulty arises from, the

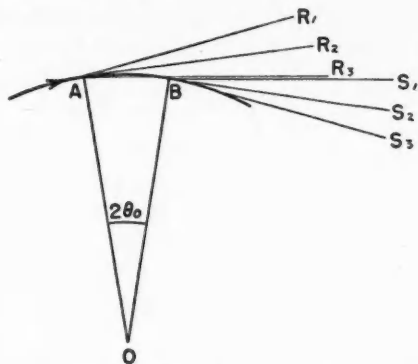


FIG. 3. A distant observer stationed along the line AR_1 or BS_1 receives radiation only when the electron is between A and B , since otherwise he lies outside the forward cone of angular opening θ_0 .

choice of a wave packet with too small an initial spread δx .

What we do instead is to minimize the final spread Δx of the packet, which is roughly $\delta x + \hbar R / mc\gamma^4 \delta x$, by choosing an optimum value for δx . It is easily seen, then, that both this optimum δx and Δx are roughly equal to $(\hbar R / mc\gamma^4)^{1/2}$. (Numerical factors of order unity are ignored in these order of magnitude estimates.) We now expect the classical theory to be valid when:

$$\begin{aligned} \Delta x &\ll \lambda_0, \quad (\hbar R / mc\gamma^4)^{1/2} \ll R/\gamma^3, \\ \gamma &\ll (mcR/\hbar)^{1/2} = (R/\lambda_0)^{1/2}. \end{aligned} \quad (4)$$

The inequality (4) is just the square root of the inequality (3), so that the two criteria are substantially equivalent although not identical; the correct criterion (4) is slightly more restrictive than the usual criterion (3). It is interesting to note that Δx turns out to be the mean proportional between the reduced de Broglie wavelength $\hbar/mc\gamma$ of the electron and the reduced wavelength λ_0 of the emitted radiation.

Dimension Perpendicular to the Plane of Motion

Let δy be the initial transverse spread of the wave packet in the direction perpendicular to the instantaneous plane of motion. Then the uncertainty principle requires a momentum spread $\delta p_y \sim \hbar/\delta y$. The spread in direction of travel of the packet due to this spread in transverse momentum is $\delta p_y/p_x$ radians, and the rate of further transverse spreading of the packet is the

product of the longitudinal speed c and this angular spread. Thus in the emission time $R/c\gamma$, the y dimension of the packet increases by

$$(\delta p_y/p_z)c(R/c\gamma) \sim \hbar R/mc\gamma^2 \delta y.$$

If now δy is chosen to minimize the final spread Δy of the packet, which is roughly $\delta y + \hbar R/mc\gamma^2 \delta y$, we find that the optimum δy and Δy are both roughly equal to $(\hbar R/mc\gamma^2)^{1/2}$. Thus, the classical theory is expected to be valid when:

$$\Delta y \ll \lambda_0/\theta_0, \quad (\hbar R/mc\gamma^2)^{1/2} \ll R/\gamma^2, \\ \gamma \ll (R/\lambda_0)^{1/2},$$

which is in agreement with Eq. (4).

Transverse Dimension in the Plane of Motion

Let δz be the initial transverse spread of the wave packet in the instantaneous plane of motion. There are two distinct effects which make the spread Δz at the end of the emission time different from δz . The first of these is just that discussed in the last paragraph, which results in a further spread of order $\hbar R/mc\gamma^2 \delta z$. The second arises from the fact that the longitudinal momentum spread $\delta p_x \sim \hbar/\delta x$ causes a spread in radius of curvature of the wave packet that is given by $\delta R = \delta p_x/p_x$, since R is proportional to p_x . Now two short circular arcs of radii R and $R + \delta R$ that start tangent to each other from the same point and have the same length s will have their end points separated by approximately $\delta(s^2/2R) \cong (s^2/2R^2)\delta R$ in the radial direction.¹¹ Since $s \sim R/\gamma$, this results in an additional

¹¹ Note added in proof: J. S. Blair has kindly pointed out to the author that this effect also introduces an additional spread into the longitudinal dimension of the wave packet that has the same order of magnitude as the terms retained. This does not therefore alter the criterion given in Eq. (4).

spread in the z direction that is roughly equal to $\hbar R/mc\gamma^2 \delta x$, where δx must be taken to be of order $(\hbar R/mc\gamma^2)^{1/2}$ in accordance with the derivation of Eq. (4). We must therefore choose δz to minimize the final spread

$$\Delta z \sim \delta z + \hbar R/mc\gamma^2 \delta z + (\hbar R/mc\gamma^2)^{1/2}.$$

It is easily seen that both the optimum δz and Δz are roughly equal to $(\hbar R/mc\gamma^2)^{1/2}$. Then as with the y dimension, the condition $\Delta z \ll \lambda_0/\theta_0$ leads immediately to Eq. (4).

It is interesting to note that the optimum wave packet has roughly the shape of an oblate spheroid, the transverse dimensions of which are γ times the longitudinal dimension. Such a packet can be formed by superposition of energy eigenfunctions (with oscillating time factors) of the electron in the magnetic field. In general, the eigenfunctions that are superposed have comparable values for their parameters, and therefore similar radiation properties. We can conclude, then, that the radiation from a representative group of energy eigenfunctions is very similar to that from our wave packet, except that in the former case the radiation pattern calculated for the packet is averaged over a revolution of the electron in the magnetic field. Then since our packet radiation is essentially the same as the classical radiation when the criterion (4) is satisfied, the same is true of the radiation from the energy eigenfunctions. It is not necessary that the wave packet continually be reformed as it spreads, or indeed that the packet exist at all. All that is necessary is that it be possible to form a wave packet with the properties stated above; when this is the case, the quantum-mechanical radiation calculation must agree with the classical calculation.

Science, as its name implies, is primarily knowledge; by convention it is knowledge of a certain kind, the kind, namely, which seeks general laws connecting a number of particular facts. Gradually, however, the aspect of science as knowledge is being thrust into the background by the aspect of science as the power of manipulating nature. It is because science gives us the power of manipulating nature that it has more social importance than art. Science as the pursuit of truth is the equal, but not the superior, of art. Science as a technique, though it may have little intrinsic value, has a practical importance to which art cannot aspire.—BERTRAND RUSSELL, The Scientific Outlook (1931).

On the Zitterbewegung of the Dirac Electron*

KERSON HUANG

Department of Physics and Laboratory for Nuclear Science and Engineering,
Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received March 3, 1952)

The detailed motion of a free Dirac electron is investigated by examining the expectation values of the position \mathbf{r} and of $\mathbf{r} \times \dot{\mathbf{r}}$ in a wave packet. It is shown that the well-known zitterbewegung may be looked upon as a circular motion about the direction of the electron spin, with a radius equal to the Compton wavelength (divided by 2π) of the electron. It is further shown that the intrinsic spin of the electron may be looked upon as the "orbital angular momentum" of this motion. The current produced by the zitterbewegung is seen to give rise to the intrinsic magnetic moment of the electron.

IT is well known that in the Dirac theory of the electron, velocity and momentum are related in an entirely different manner from that in non-relativistic mechanics. While in the latter case the velocity operator is simply proportional to the momentum operator, in the Dirac theory it is represented by the operator α , whose components can only have eigenvalues of $\pm c$, where c is the velocity of light. This means that while the average velocity of the electron is less than c , the instantaneous velocity is always $\pm c$. We infer from this that the motion of the electron consists of a highly oscillatory component, superimposing on the average motion. Schrödinger¹ called this oscillatory motion "zitterbewegung," and showed that the amplitude of the oscillation is of the order of the Compton wavelength of the electron.

Nevertheless, the zitterbewegung is not an observable motion, for any attempt to determine the position of the electron to better than a Compton wavelength must defeat its purpose by the creation of electron-positron pairs. This may be stated in a more formal way by saying that the position operator in the Dirac theory is not a "one-particle observable," in the sense that if ψ represents the wave function of the electron in a state having purely positive or negative energy eigenvalues, the state $\mathbf{x}\psi$ will fail to have the same property. Thus, the trajectory of the electron can never be traced out in a well-defined path. It can at best be confined to a "tube" whose diameter is of the order of a Compton wavelength.

It is hence meaningless from a physical viewpoint to inquire about the detailed motion in the "tube." However, it is not improper to ask whether we can form some picture of this detailed motion, be it a rough picture which we must use with reservation.

Our desire to have some picture of the zitterbewegung is further motivated by the intimate connection between zitterbewegung and the intrinsic magnetic moment (and hence intrinsic spin) of the electron. Gordon² has shown that the current density of a free electron can be brought to the form:

$$\mathbf{j} = \mu_0 [(\mathbf{grad} \bar{\psi})\psi - \bar{\psi}(\mathbf{grad} \psi)] + \mathbf{curl} \mathbf{m} + \frac{1}{c} \frac{\partial \mathbf{P}}{\partial t}, \quad (1)$$

where $\bar{\psi}$ is the adjoint of the spinor wave function ψ , (related to the Hermitian adjoint ψ^* by $\bar{\psi} = i\psi^* \beta$), $\mu_0 = e\hbar/2me$ where e is the algebraic charge of the electron, and

$$\mathbf{m} = \mu_0 \psi^* \beta \boldsymbol{\sigma} \psi, \quad \mathbf{P} = \mu_0 \psi^* \boldsymbol{\gamma} \psi, \quad (2)$$

β , $\boldsymbol{\sigma}$, and $\boldsymbol{\gamma}$ being the usual Dirac matrices. The first term in \mathbf{j} is just the familiar expression for the current density of a spinless particle, while the remaining two terms are interpreted to be contributions coming from a magnetic moment \mathbf{m} and an electric moment \mathbf{P} , respectively. The contribution due to \mathbf{m} is particularly interesting, for its form,

$$\mathbf{j}' = \mathbf{curl} \mathbf{m} \quad (3)$$

can be shown to be equivalent to the relation

$$\mathbf{m} = \frac{1}{2} \mathbf{r} \times \mathbf{j}' \quad (4)$$

* Supported in part by the joint program of the ONR and the AEC.

¹ E. Schrödinger, Sitzber. preuss. Akad. Wiss., Physik-math. Kl. 24, 418 (1930).

² W. Gordon, Z. Physik 50, 630 (1927); J. Frenkel, *Wave Mechanics, Advanced General Theory* (Oxford University Press, London, 1934), p. 321.

if one performs a space average³ i.e.,

$$\int \mathbf{m} dV = \frac{1}{2} \int \mathbf{r} \times \text{curl } \mathbf{m} dV.$$

Equation (4), which is the classical expression for the magnetic moment of a loop of current, suggests that \mathbf{m} has its origin in the zitterbewegung. It is very tempting to say that the magnetic moment and the spin of the electron arise from some sort of circular motion of the electron. We shall show that this picture is not altogether an unreasonable one.

Let us examine how the position \mathbf{r} of the electron behaves on the average in a state in which the electron is localized in space. For this purpose, we shall construct a wave packet for the electron, and find the expectation value of \mathbf{r} . This may be done easily by first finding $\langle \mathbf{r} \rangle = c \langle \boldsymbol{\alpha} \rangle$ and then find $\langle \mathbf{r} \rangle$ by an integration.

The general Dirac wave packet may be written in a momentum space expansion of the form

$$\psi(\mathbf{r}, t) = \hbar^{-1} \int [C^+(\mathbf{p})e^{-i\omega t} + C^-(\mathbf{p})e^{i\omega t}] \exp(i\mathbf{p} \cdot \mathbf{r}/\hbar) d^3\mathbf{p}, \quad (5)$$

where $\omega = \epsilon/\hbar$, $\epsilon = \pm(c^2p^2 + m^2c^4)^{1/2}$, $C^+(\mathbf{p})$ is a linear combination of "spin up" and "spin down" amplitudes of the free particle Dirac waves belonging to the momentum \mathbf{p} with positive energy, and $C^-(\mathbf{p})$ for negative energy. That is,

$$\begin{aligned} C^+(\mathbf{p}) &= a_1 u_1 + a_2 u_2 \\ C^-(\mathbf{p}) &= a_3 u_3 + a_4 u_4, \end{aligned} \quad (6)$$

where the coefficients a_k are complex functions of \mathbf{p} , and

$$\begin{aligned} u_1 &= \begin{bmatrix} 1 \\ 0 \\ kp_3 \\ kp_+ \end{bmatrix}, & u_2 &= \begin{bmatrix} 0 \\ 1 \\ kp_- \\ -kp_3 \end{bmatrix}, \\ u_3 &= \begin{bmatrix} -kp_3 \\ -kp_+ \\ 1 \\ 0 \end{bmatrix}, & u_4 &= \begin{bmatrix} -kp_- \\ kp_3 \\ 0 \\ 1 \end{bmatrix}, \end{aligned} \quad (7)$$

with

$$k = \frac{c}{\epsilon + mc^2}, \quad p_{\pm} = p_1 \pm ip_2. \quad (8)$$

The wave packet, Eq. (5), is normalized provided we require that

$$\int [C^{+*}(\mathbf{p})C^+(\mathbf{p}) + C^{-*}(\mathbf{p})C^-(\mathbf{p})] d^3\mathbf{p} = 1.$$

Hence,

$$\begin{aligned} \langle \mathbf{r} \rangle &= c \langle \boldsymbol{\alpha} \rangle = c \int (C^{+*} \boldsymbol{\alpha} C^+ + C^{-*} \boldsymbol{\alpha} C^-) d^3\mathbf{p} \\ &+ c \int [(C^{+*} \boldsymbol{\alpha} C^-) e^{2i\omega t} + (C^{-*} \boldsymbol{\alpha} C^+) e^{-2i\omega t}] d^3\mathbf{p}, \end{aligned}$$

the arguments of C^{\pm} being understood. Since $\boldsymbol{\alpha}$ is a Hermitian operator, we have

$$\begin{aligned} \langle \mathbf{r} \rangle &= c \int (C^{+*} \boldsymbol{\alpha} C^+ + C^{-*} \boldsymbol{\alpha} C^-) d^3\mathbf{p} \\ &+ 2c \int \text{Re}\{(C^{+*} \boldsymbol{\alpha} C^-) e^{2i\omega t}\} d^3\mathbf{p}. \end{aligned}$$

By making use of the relations⁴

$$C^{\pm*} \boldsymbol{\alpha} C^{\pm} = \pm \frac{c\mathbf{p}}{\epsilon} (C^{\pm*} C^{\pm}),$$

the above may be rewritten

$$\langle \mathbf{r} \rangle = \bar{\mathbf{v}} + 2c \int \mathbf{K}(\mathbf{p}) \sin(2\omega t + \phi(\mathbf{p})) d^3\mathbf{p}, \quad (9)$$

where

$$\begin{aligned} \bar{\mathbf{v}} &= c \int \frac{\mathbf{p}}{mc} \left[1 + \left(\frac{\mathbf{p}}{mc} \right)^2 \right]^{-1/2} (C^{+*} C^+ - C^{-*} C^-) d^3\mathbf{p}, \\ \mathbf{K}(\mathbf{p}) &= |C^{-*} \boldsymbol{\alpha} C^+|, \end{aligned}$$

$$\phi(\mathbf{p}) = \tan^{-1} \{ \text{Re}(C^{-*} \boldsymbol{\alpha} C^+) / \text{Im}(C^{-*} \boldsymbol{\alpha} C^+) \}. \quad (10)$$

Hence

$$\begin{aligned} \langle \mathbf{r} \rangle &= \mathbf{r}^0 + \bar{\mathbf{v}} t - \lambda \int \mathbf{K}(\mathbf{p}) [1 + (\mathbf{p}/mc)^2]^{-1/2} \\ &\times \cos(2\omega t + \phi(\mathbf{p})) d^3\mathbf{p}, \end{aligned} \quad (11)$$

⁴ See W. Pauli, *Handbuch der Physik* (J. Springer, Berlin, 1933), 2 Aufl., Band 24, 1. Teil, p. 228.

³ See reference 2, Frenkel.

where $\lambda = \hbar/mc$ is the Compton wavelength (divided by 2π) of the electron.

The expectation value for the position is seen to be composed of two parts: a time-independent part, which describes the average motion of the electron, and a time-dependent part, which involves interference terms of the type $C^+ \alpha C^-$. It arises from the interference between the positive and negative energy states of the electron, and represents the zitterbewegung of the electron.

The results obtained so far hold for any state ψ of the electron. To investigate the properties of the time-dependent part of $\langle \mathbf{r} \rangle$, it is necessary to have an explicit form of ψ in order to know the exact forms of $C^\pm(\mathbf{p})$. For this purpose, we shall construct a wave packet which satisfies the initial condition:

$$\psi(\mathbf{r}, 0) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} f(\mathbf{r}/r_0), \quad (12)$$

where $f(\mathbf{r}/r_0)$ is a normalized Gaussian function:

$$f(\mathbf{r}/r_0) = (2/\pi r_0^2)^{1/2} \exp[-(\mathbf{r}/r_0)^2], \quad (13)$$

where $\mathbf{r} = |\mathbf{r}|$, and r_0 is a constant, giving the approximate extension of the wave packet in space. Thus initially the wave packet has its spin pointing along the x_3 -axis, and the center of the packet is at rest at the origin. With this initial condition, $\psi(\mathbf{r}, t)$ is completely specified, and we can calculate for C^\pm , giving

$$C^+(\mathbf{p}) = \frac{f(\mathbf{p}/\eta)}{1+k^2 p^2} \begin{bmatrix} 1 \\ 0 \\ k p_3 \\ k p_+ \end{bmatrix},$$

$$C^-(\mathbf{p}) = \frac{f(\mathbf{p}/\eta)}{1+k^2 p^2} \begin{bmatrix} k^2 p^2 \\ 0 \\ -k p_3 \\ -k p_+ \end{bmatrix}, \quad (14)$$

where $f(\mathbf{p}/\eta)$ is the Fourier transform of $f(\mathbf{r}/r_0)$:

$$f(\mathbf{p}/\eta) = (2/\pi \eta^2)^{1/2} \exp[-(\mathbf{p}/\eta)^2], \quad (15)$$

satisfying the normalizing condition

$$\int f^2(\mathbf{p}/\eta) d^3 \mathbf{p} = 1,$$

where η is a constant giving the band width of the wave packet in momentum space. It is related to r_0 by

$$\eta = 2\hbar/r_0. \quad (16)$$

The wave packet has the appearance:

$$\psi(\mathbf{r}, t) = \hbar^{-1} \int \begin{bmatrix} 1 \\ 0 \\ k p_3 \\ k p_+ \end{bmatrix} e^{-i\omega t} + \begin{bmatrix} k^2 p^2 \\ 0 \\ -k p_3 \\ -k p_+ \end{bmatrix} e^{i\omega t} \frac{f(\mathbf{p}/\eta)}{1+k^2 p^2} \exp(i\mathbf{p} \cdot \mathbf{r}/\hbar) d^3 \mathbf{p}. \quad (17)$$

In order that the wave packet does not spread out too rapidly in time, we choose the constant r_0 very large compared to a Compton wavelength. This means that η is very small compared to mc :

$$\eta/mc \ll 1.$$

Then in the integrand in Eq. (17), all the p 's greater than η will be attenuated and essentially make no contributions to the integral, due to the presence of the factor $\exp[-(\mathbf{p}/\eta)^2]$ in $f(\mathbf{p}/\eta)$. This means that for all the momenta that contribute to the integral we have $p/mc \ll 1$, and hence we may use the nonrelativistic approximation $\epsilon \cong mc^2$, $k \cong \frac{1}{2}mc$. This being the case, one can see that in Eq. (17) C^- is smaller than C^+ by one order in p/mc (the "large" components of C^+ are the upper two components while that of C^- are the lower ones). Furthermore, in both C^+ and C^- the "spin up" components are larger than the "spin down" ones by one order in p/mc . This means that the wave packet is built up of mainly positive energy waves, and that all the Fourier waves in the decomposition have their spins pointing approximately in the x_3 -direction. It should be noted that the negative energy components appear for two reasons, firstly because we have localized the electron, and secondly because we have limited the direction of the spin, both \mathbf{r} and σ being not "one particle observables."

For the wave packet, Eq. (17), the time independent terms in $\langle \mathbf{r} \rangle$ vanishes. We are left only with the zitterbewegung. This can now be evaluated, making the approximation that $\epsilon \cong mc^2$, which amounts to neglecting terms of the order of $(p/mc)^2$. The results are:

$$K_1(\mathbf{p}) = K_2(\mathbf{p}) = \frac{f^2(\mathbf{p}/\eta)}{2mc} (p_1^2 + p_2^2)^{1/2} \quad (18)$$

$$K_3(\mathbf{p}) = \frac{f^2(\mathbf{p}/\eta)}{2mc} p_3,$$

and

$$\tan \varphi_1 = -p_1/p_2, \quad \tan \varphi_2 = p_2/p_1, \quad \tan \varphi_3 = \infty. \quad (19)$$

It can be immediately seen that

$$\varphi_1 = \varphi_2 + \pi, \quad \varphi_3 = \pi/2.$$

For convenience in evaluating $\langle \mathbf{r} \rangle$ we shall introduce spherical coordinates in momentum space, so that each plane wave in the Fourier decomposition is characterized by the spherical coordinates of its momentum vector $\mathbf{p}(p, \vartheta, \varphi)$, such that

$$\begin{aligned} p_1 &= p \sin \vartheta \cos \varphi \\ p_2 &= p \sin \vartheta \sin \varphi \\ p_3 &= p \cos \vartheta. \end{aligned}$$

We find that in this coordinate system

$$\begin{cases} K_1 = K_2 = \frac{f^2(p/\eta)}{2mc} p \sin \vartheta \\ K_3 = \frac{f^2(p/\eta)}{2mc} p \cos \vartheta \\ \varphi_1 = \varphi + \pi/2 \\ \varphi_2 = \varphi, \quad \varphi_3 = \pi/2. \end{cases} \quad (20)$$

Hence

$$\begin{aligned} \langle x_1 \rangle &= -I \int_0^{2\pi} \lambda \sin(2\omega t + \varphi) d\varphi \\ \langle x_2 \rangle &= -I \int_0^{2\pi} \lambda \cos(2\omega t + \varphi) d\varphi \\ \langle x_3 \rangle &= -\pi \lambda \int_0^\infty \frac{f^2(p/\eta)}{2mc} p^2 dp \\ &\quad \times \int_0^\pi \sin 2\vartheta d\vartheta \sin 2\omega t, \end{aligned} \quad (21)$$

where

$$\begin{aligned} I &= - \int_0^\infty \frac{f^2(p/\eta)}{2mc} p^3 dp \int_0^\pi \sin^2 \vartheta d\vartheta \\ &= -(32\pi)^{-1} (\lambda/r_0). \end{aligned}$$

Although all components of $\langle \mathbf{r} \rangle$ vanish upon integration, it is interesting to note the forms of $\langle x_1 \rangle$ and $\langle x_2 \rangle$. If we consider those Fourier waves which have momentum vectors making an azimuthal angle between φ and $\varphi + d\varphi$, we

may say that their contribution to the motion is

$$\begin{aligned} \langle x_1 \rangle_\varphi &= -\lambda \sin(2\omega t + \varphi) \\ \langle x_2 \rangle_\varphi &= -\lambda \cos(2\omega t + \varphi). \end{aligned} \quad (22)$$

The factor $I \sim \lambda/r_0$ in Eq. (21) comes from radial and ϑ integrations, and it only determines the weight of zitterbewegung in the total motion of the electron, relative to the degree of localization of the electron in space. Equation (22) shows that each plane wave in the Fourier decomposition contributes a motion which is circular in the x_1 - x_2 plane, with radius λ . Since this motion is independent of the p, ϑ coordinates of the momentum, and since by a previous argument we know that the spin of each plane wave is approximately in the x_3 -direction, we may say that the contribution of each plane wave to the motion of the wave packet is an orbit whose projection on the plane perpendicular to the spin is a circle of radius λ , the phase of the motion in the circle being equal to φ . The total contribution from all the plane waves vanishes, because the wave packet is spherically symmetric, and for each circular motion with phase φ , there is a corresponding motion with phase $\varphi + \pi$ to cancel it. Hence, the center of the wave packet as a whole remains at rest.

To gain further insight into the situation, we shall examine the expectation value of the operator $\mathbf{r} \times \hat{\mathbf{f}}$:

$$\langle \mathbf{r} \times \hat{\mathbf{f}} \rangle = c \langle \mathbf{r} \times \boldsymbol{\alpha} \rangle - i\hbar c \langle \nabla_p \times \boldsymbol{\alpha} \rangle, \quad (23)$$

where $(\nabla_p)_k = \partial/\partial p_k$. After some straightforward calculations we find, again to the approximation in which $(p/mc)^2$ is neglected:

$$\begin{aligned} \langle \mathbf{r} \times \hat{\mathbf{f}} \rangle_1 &= 0, \quad \langle \mathbf{r} \times \hat{\mathbf{f}} \rangle_2 = 0 \\ \langle \mathbf{r} \times \hat{\mathbf{f}} \rangle_3 &= (\hbar/m)(1 - \cos 2\omega t). \end{aligned} \quad (24)$$

It is interesting to point out that $\langle \mathbf{r} \times \hat{\mathbf{f}} \rangle_1$ and $\langle \mathbf{r} \times \hat{\mathbf{f}} \rangle_2$ not only vanish on the average, but the contribution from each Fourier component vanishes individually. Since classically and for circular motion $\mathbf{r} \times \hat{\mathbf{f}}$ is a vector perpendicular to the plane of motion, this result lends support to our picture in which each Fourier wave contributes a circular motion about the direction of $\mathbf{r} \times \hat{\mathbf{f}}$, or the direction of the spin. That this is so follows from the fact that $\mathbf{r} \times \hat{\mathbf{f}}$ is proportional to the magnetic moment of the electron. As a

matter of fact the magnetic moment

$$\mathbf{u} = (1/2c)\mathbf{r} \times \mathbf{j} = (e/2c)\mathbf{r} \times \mathbf{f}$$

has the expectation values

$$\langle \mu_1 \rangle = 0, \quad \langle \mu_2 \rangle = 0 \\ \langle \mu_3 \rangle = (e\hbar/2mc)(1 - \cos 2\omega t), \quad (25)$$

with the correct g factor.

Thus, we may picture the intrinsic magnetic moment to be a result of the zitterbewegung. We can also imagine the intrinsic spin to arise in the same manner, due to the intimate connection between magnetic moment and spin.

We can further justify the picture that both intrinsic magnetic moment and spin arise from the zitterbewegung by calling attention to the following identity, which may be established by the method of the "Gordon split":²

$$mc\psi^* \mathbf{r} \times \boldsymbol{\alpha} \psi = \psi^* \boldsymbol{\beta} \mathbf{r} \times \mathbf{p} \psi + \hbar \psi^* \boldsymbol{\beta} \boldsymbol{\sigma} \psi \\ + \frac{\hbar}{2ic} \frac{\partial}{\partial t} \{ \psi^* \boldsymbol{\beta} (\mathbf{r} \times \boldsymbol{\alpha}) \psi \}. \quad (26)$$

The last term on the right vanishes if ψ^*, ψ both refer to the same state. Hence, we can rewrite it as follows:

$$\hbar \bar{\psi} \boldsymbol{\sigma} \psi = \bar{\psi} \mathbf{r} \times (m\boldsymbol{\beta} \mathbf{f} - \mathbf{p}) \psi. \quad (27)$$

Recalling that $\langle \beta \rangle = \langle \beta^{-1} \rangle = [1 - (v/c)^2]^{-1/2}$, we see that $m\boldsymbol{\beta} = m\beta^{-1}$ represents the relativistic mass, and hence in a sense, the spin is an "orbital angular momentum" due to the difference between instantaneous velocity \mathbf{f} and average velocity \mathbf{p}/m —or, zitterbewegung. Rewriting Eq. (27) in another form and introducing the abbreviation $\mathbf{u} = (e/2c)(\mathbf{r} \times \mathbf{f})$, we have

$$\psi^* \mathbf{u} \psi = (e/2mc)\psi^* \boldsymbol{\beta} \mathbf{L} \psi + (e/mc)\psi^* \boldsymbol{\beta} \mathbf{S} \psi, \quad (28)$$

where $\mathbf{L} = \hbar \mathbf{r} \times \mathbf{p}$ is the orbital angular momentum and $\mathbf{S} = (\hbar/2)\boldsymbol{\sigma}$ the spin angular momentum, with the factors β again reflecting the variation of mass with velocity. The above expression shows that $(e/2c)\mathbf{r} \times \mathbf{f}$ represents the total magnetic moment of the electron, both orbital and intrinsic, with the correct g factors for both cases: $g_L = \hbar/2mc$, $g_S = \hbar/mc$. It is interesting to note that while $\mathbf{r} \times \mathbf{f}$ takes into account the total magnetic moment of the electron, there is no simple relationship between $\mathbf{r} \times \mathbf{f}$ and the total angular momentum $\mathbf{r} \times \mathbf{p} + (\hbar/2)\boldsymbol{\sigma}$. This merely

stresses again the basic difference between velocity and momentum in relativity.

It may not be superfluous to point out that the expression (24) for $\langle \mu_3 \rangle$ contains a time-dependent term (interference between positive and negative energy states) regardless of the "diffusedness" of the wave packet in space. This is, of course, due to the fact that in our wave packet, we have not only localized the electron, but also demanded that it has a well-defined initial spin direction. This last requirement amounts to an exact measurement of spin at the initial time. The measurement disturbs the electron state and brings in negative energy components which cannot be removed by simply letting the packet spread in space. To verify this, we may impose a different initial condition in which no attempt is made to observe the spin accurately. We may, for example, make the initial packet an eigenfunction of the "mean spin operator" Σ of Foldy and Wouthuysen,⁵ where Σ is, roughly speaking, the "one particle part" of σ :

$$\Sigma = \sigma - (i/\epsilon)\boldsymbol{\beta} \boldsymbol{\alpha} \times \mathbf{p} - [\epsilon(\epsilon + mc^2)]^{-1} \mathbf{p} \times (\boldsymbol{\sigma} \times \mathbf{p}). \quad (29)$$

Our original initial packet can be made an eigenfunction of Σ , by applying to it the projection operator $(\Sigma_3 + 1)/2$, yielding

$$\psi(\mathbf{r}, 0) = \frac{1}{2}(\Sigma_3 + 1) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} f(r/r_0) = \begin{pmatrix} 1 - (k/2\epsilon)p_+p_- \\ (k/2\epsilon)p_+p_3 \\ 0 \\ (k/2\epsilon)p_+ \end{pmatrix} f(r/r_0).$$

Repeating the calculations with such a wave packet, we find that

$$\langle \mu_1 \rangle = \langle \mu_2 \rangle = 0, \quad \langle \mu_3 \rangle = e\hbar/2mc.$$

In conclusion, it may be mentioned that many "classical models" of the Dirac electron have been constructed by different workers basing on the zitterbewegung. Hönl,⁶ for example, considers the electron to be a point charge moving with the velocity of light in a circle whose radius is $\hbar/2mc$. The spin has the correct value $\hbar/2$, but

⁵ L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).

⁶ H. Hönl, Ann. Physik (V) **33**, 565 (1938).

the magnetic moment has to be assumed phenomenologically. It has even been suggested that the electron has further a "hyperzitterbewegung," some sort of epicyclic motion, which accounts for the Schwinger correction to the magnetic moment.⁷ Mathissen⁸ and Weyssenhof⁹ proposed similar models basing their picture on

⁷ J. Brandmüller, *Naturwiss.* **38**, 139 (1951).

⁸ M. Mathissen, *Acta Phys. Polon.* **6**, 163 (1937); 218 (1937).

⁹ J. W. Weyssenhof, *Nature* **141**, 328 (1938); *Acta Phys. Polon.* **9**, 1 (1947).

the helical path taken by a spinning particle in general relativity. We are, however, not concerned with classical models here. The aim of this note is to acquire from the Dirac equation itself detailed information concerning the motion of an electron.

The author wishes to express his deep gratitude to Professor V. F. Weisskopf, under whose guidance this investigation was carried out. Thanks are also due to Dr. F. Villars for helpful discussions.

American Physicists and Their Graduate Degrees*

JOHN N. COOPER

The Ohio State University, Columbus 10, Ohio

(Received April 4, 1952)

There are some 4200 physicists listed in the 1949 edition of *American Men of Science*. Roughly three-fourths have a doctoral degree; only about eight percent have no graduate degree. The number of graduate degrees granted to this group by each of the major American graduate schools has been tabulated. More than 1800 of the 3000 doctoral degrees were awarded by thirteen leading institutions, each of which granted over one hundred Ph.D. degrees. The University of Chicago has awarded the largest number, 230. After Chicago come Cornell, California (Berkeley), Michigan, California Institute of Technology, Johns Hopkins, Harvard, Massachusetts Institute of Technology, Columbia, Yale, Wisconsin, Princeton, and Illinois. Over sixty other North American schools and fifty European institutions have granted one or more of the Ph.D. degrees.

A survey of the doctorates awarded during successive five-year periods reveals significant changes in the order of productivity of institutions. Within the past decade many schools have expanded their graduate programs to the point at which they are making significant contributions in training at the doctoral level.

THE training of professional physicists is a major responsibility of American graduate schools. Not only is the profession dependent on a suitable influx of young physicists, but national security demands it as well. In recent years considerable interest has been shown in statistics on the production of Ph.D.'s. In 1941 Birge¹ reported on the training and prospects of the Ph.D. in physics. Four years later Trytten² accurately predicted an impending scarcity of scientific personnel as a result of the shunting of promising young men from graduate schools to the war

effort. More recently White³⁻⁵ has made several reports on the number of physicists in training, both at the doctoral and at lower levels. Murdoch and White discussed the places of birth⁶ and the geographic region in which physicists earned their Ph.D.'s, as well as the occupations⁷ in which physicists who earned their Ph.D.'s between 1936 and 1948 found themselves in 1949-1950. Bell⁸ tabulated both the graduate and undergraduate origins of all physicists listed in the 1944 edition of *American Men of Science*. The present paper

³ M. W. White, *Rev. Sci. Instr.* **18**, 219 (1947).

⁴ M. W. White, *Phys. Today* **2**, No. 11, 16 (1949).

⁵ M. W. White, *Am. J. Phys.* **19**, 27 (1951).

⁶ B. C. Murdoch and M. W. White, *Phys. Today* **3**, No. 12, 17 (1950).

⁷ B. C. Murdoch and M. W. White, *Phys. Today* **3**, No. 9, 20 (1950).

⁸ R. M. Bell, *Am. J. Phys.* **14**, 396 (1946).

* Based upon a paper presented at the meeting of the American Association of Physics Teachers in New York, January 31, 1952.

¹ R. T. Birge, *Am. J. Phys.* **9**, 24 (1941).

² M. H. Trytten, *Sci. Monthly* **60**, 37 (1945).

represents a comparable survey of the graduate school origins of physicists listed in the most recent edition of this book.

The names of more than 4200 physicists are presented in the 1949 edition of *American Men of Science*. Slightly fewer than three-fourths of them have doctoral degrees; of those who do not, roughly two-thirds have a master's degree and the remaining third have only the bachelor's degree. The number of advanced degrees, both master's and doctor's, granted by each of the major American graduate schools has been ascertained. The sole criterion used in determining whether or not a person should be counted was that the word *physics* appear in the field of specialization. Thus, men in astrophysics and

TABLE II. The number of doctor's and master's degrees awarded to physicists according to the 1949 edition of *American Men of Science*.

Rank	Institution	Number of degrees awarded		Rank	Institution	Number of degrees awarded	
		Doctor's	Master's			Doctor's	Master's
14	Ohio State	73	82	24	Toronto	32	49
15	New York	65	48	25	Pittsburgh	30	35
16	Iowa	62	78	26	Purdue	29	17
17	Minnesota	59	40	27	McGill	27	27
18	Pennsylvania	58	30	28	Brown	25	27
19	Indiana	53	55	29	Texas	21	35
20	Virginia	46	48	30	Cincinnati	21	20
21	Pennsylvania State	41	35	31	Duke	20	15
22	(Rochester)	33	17		Iowa State	20	25
	(Stanford)	33	13		Washington (Seattle)	20	21

TABLE I. The number of doctor's and master's degrees awarded to physicists in the 1949 edition of *American Men of Science* by the thirteen American graduate schools which have awarded more than one hundred doctorates. The total number of doctorates is broken down into the number granted in each of seven time intervals.

Rank	Institution	Doctor's degrees awarded								No. of master's degrees
		Pre '21	21-25	26-30	31-35	36-40	41-45	46-49	Total	
1	Chicago	40	31	35	33	41	28	22	230	85
2	Cornell	41	19	26	24	25	29	15	179	45
3	California (Berkeley)	9	8	25	33	42	34	20	171	82
4	Michigan	8	7	26	30	42	21	16	150	142
5	California Institute of Technology	—	10	35	37	25	22	19	148	49
6	Johns Hopkins	45	19	19	17	20	12	12	144	26
7	Harvard	24	8	14	31	22	15	27	141	129
8	Massachusetts Institute of Technology	1	15	4	14	39	35	41	139	68
9	Columbia	19	11	16	23	26	15	14	124	125
10	Yale	17	13	12	23	22	14	20	121	28
11	Wisconsin	12	11	18	26	22	14	13	116	103
12	Princeton	17	11	17	15	24	12	13	109	50
13	Illinois	12	4	14	18	22	26	6	102	117

biophysics were counted, while those in astronomy and engineering were not. Some scientists with degrees in physics have thus been missed, while others with degrees in related fields have been included if their present field of specialization includes physics. The survey by Murdoch and White⁷ shows that 80 percent of the 1700 physicists who earned a doctorate between 1936 and 1948 are still active in physics; part of the remaining 20 percent are doing educational or administrative work and would doubtless continue to list themselves as physicists. This suggests that the number of misassignments should be less than 15 percent of the total number.

Of the listed 3000 physicists with doctoral degrees the vast majority have Ph.D.'s; only

3 percent have earned D.Sc. degrees. Well over 1800, or 60 percent, of the doctorates were bestowed by thirteen leading institutions, each of which granted over one hundred doctorates. The numbers awarded by these schools are shown in Table I. The total numbers are broken down into the numbers of awards made before 1921 and the numbers granted in successive five-year periods after 1920. The University of Chicago leads in total number of Ph.D.'s awarded by a large margin. Of the thirteen institutions in Table I only Massachusetts Institute of Technology has awarded a significant number of earned D.Sc. degrees, a total of thirty.

A survey of the degrees granted during successive five-year periods reveals that significant changes in the list of most productive institutions have occurred. Outstanding is the rise of California, Michigan, California Institute of Technology, and Massachusetts Institute of Technology, none of which had awarded many doctoral degrees before 1920. On the other hand, Johns Hopkins University, first in the number of doctorates granted before 1920, ranks last of the thirteen institutions in degrees awarded since 1940. Comparison of Table I with the corresponding table in Bell's compilation⁸ from the 1944 edition of *American Men of Science* shows the same thirteen institutions in the lead, but with several major changes in order.

Table II shows all other American institutions which have granted twenty or more doctorates to persons listed in *American Men of Science*. The word "American" is used here in the sense of North American, as is evident from the fact that the table includes the names of two Canadian institutions. These twenty institutions contrib-

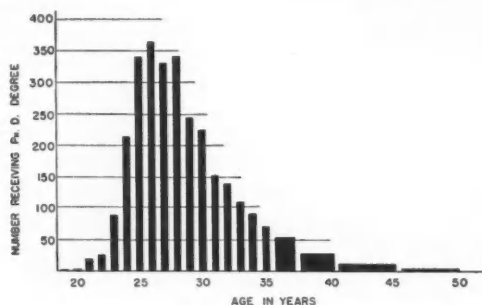


FIG. 1. The number of physicists listed in the 1949 edition of *American Men of Science* who received their Ph.D.'s at a given age.

uted a total of 768 doctorates, approximately 25 percent of the total. Thus, over 85 percent of the doctorates were awarded by the 33 institutions which granted twenty or more.

Close behind the universities of Table II come George Washington University, Catholic University, Rice Institute, Clark, Missouri, North Carolina, Northwestern, St. Louis, Washington University (St. Louis), Colorado, and Notre Dame; each has awarded ten or more doctorates. All in all, some 75 American schools and over 50 European institutions have awarded doctorates to physicists listed in *American Men of Science*. Of the universities abroad Cambridge leads in degrees awarded to physicists with 26 Ph.D.'s.

Tables I and II show not only the doctorates awarded by the listed institutions, but also the master's degrees. The number of master's degrees is not readily correlated with the number of doctorates for a number of reasons. First, some institutions automatically award a master's degree upon the satisfactory completion of a given number of hours of work, while others require a thesis. In institutions which require master's theses, men are often encouraged to work directly toward a doctor's degree. Another factor which may be important has to do with the degree of selection which various graduate schools exercise in admitting candidates for advanced degrees. Some schools try to select only persons who are of Ph.D. caliber and use the master's degree as a device for stopping those who do not seem to be promising physicists, while at the other extreme are institutions which

regularly encourage their better men to take the master's degree.

The number of persons to receive the Ph.D. at a given age is shown in Fig. 1 as a function of age. The strong maximum near 26 suggests that the majority continue toward the advanced degrees without extended periods of work outside of academic institutions. However, a substantial number of physicists have gained considerable experience before proceeding toward the doctorate.

The general impression of Fig. 1 that the majority of persons proceed directly toward advanced degrees is confirmed by the data of Table III which relate to the time elapsed between the bachelor's and the Ph.D. degrees. In each of the time intervals of Table III the largest number of physicists completed work for the doctorate four years after receiving the bachelor's degree, except that after 1945 the greatest single group received degrees after ten years, the next largest single group after six years, and the third largest after seven years. Over 20 percent of the persons receiving degrees after 1946 had earned the bachelor's degree more than ten years before they completed their doctoral requirements; a similar pattern of slightly smaller extent is indicated after World War I. This presumably reflects the fact that during the war years many young men were active in the armed forces or on projects connected with the war efforts; thus, they were delayed in completing their graduate work.

The vast majority of physicists have taken work at more than one institution. Of those

TABLE III. The time expired between bachelor's and Ph.D. degrees for physicists listed in the 1949 edition of *American Men of Science*.

Time in years	Number receiving Ph.D. degrees in							Total
	Pre 21	21- 25	26- 30	31- 35	36- 40	41- 45	46- 48	
1	0	0	1	1	0	0	0	2
2	6	5	3	7	3	4	0	28
3	7	29	54	56	59	54	11	310
4	50	35	68	116	138	90	20	517
5	39	19	54	79	114	100	35	440
6	42	24	45	72	88	59	60	390
7	30	17	37	45	60	33	53	275
8	22	16	21	39	41	27	36	202
9	16	15	22	19	26	15	43	156
10	25	11	12	24	22	13	23	130
10+	46	41	59	74	87	49	72	428
Total	323	212	376	532	638	444	353	2878

with a Ph.D. 435 (less than 15 percent) took bachelor's, master's, and Ph.D. degrees at the same institution, while 268 (9 percent) took bachelor's and Ph.D.'s at the same institution and avoided the master's degree. Only thirteen men took bachelor's and doctorates at the same institution, but went elsewhere for a master's degree. Some 451 (15 percent) took bachelor's and master's degrees at the same institution but Ph.D.'s elsewhere, while 674 (22 percent) took master's and Ph.D.'s at the same institution after previously earning the bachelor's degree elsewhere.

During the past fifteen years a number of institutions have begun to offer work at the Ph.D. level and several have already trained five or more Ph.D.'s. These institutions are currently making contributions which are of great signifi-

cance, although they do not appear in any of the tables herein. Some of these schools show promise of becoming major contributors in the training of graduate students and within a relatively short span may find themselves in the top 25 graduate schools from the point of view of Ph.D.'s awarded in physics. White's surveys^{4,5} of graduate school enrollments in 1948-1949 and in 1949-1950 reveal that many institutions not listed in Table I are now among the first dozen schools in graduate enrollment in physics. Institutions which have tremendously expanded their graduate programs include Maryland, New York University, California (Los Angeles), Cincinnati, Polytechnic Institute of Brooklyn, and Catholic University. The effects of this expansion should become evident within a decade or two.

Free Tape Recordings of Speeches

of

E. U. CONDON, K. K. DARROW, AND E. FERMI

Through the efforts of Dr. A. R. Tobey, Supervisor, Electricity and Magnetism Section, Armour Research Foundation, and with permission of the American Institute of Physics, free tape recordings are available on loan to members of the Association who wish to use them for presentation to classes, seminars, or science clubs. The subjects are three of the six lectures presented as a Symposium on Physics Today at the Twentieth Anniversary meeting of the American Institute of Physics, Chicago, Illinois, on October 25, 1951, as follows:

The Atom. E. U. CONDON.

Physics as Science and Art. K. K. DARROW.

The Nucleus. E. FERMI.

Each recording is complete on one reel, two tracks. It must be played on a double-track tape machine at $3\frac{3}{4}$ inches per sec. No rewinding is necessary. Track No. 1 should be played to the

end, and then the reels should be reversed (and turned over) to play track No. 2. This procedure rewinds the reel into its original condition.

Requests for recordings should be made to THOMAS H. OSGOOD, Editor, *American Journal of Physics*, Michigan State College, East Lansing, Michigan. To allow some flexibility in scheduling, a request should specify a preferred date and two alternative acceptable dates.

No charge is made for the loan of recordings. The borrower is expected, however, to pay the return postage to the Editor's office or to any other address the Editor specifies. Transportation is to be by first class mail, special delivery, which will cost less than one dollar per reel at present rates. Reels will be mailed in convenient reusable boxes.

Prospective borrowers may consult the published papers in *Physics Today*, November, 1951, January, 1952, and March, 1952.

The Concept of Temperature

A. R. MILLER

Cavendish Laboratory, University of Cambridge, Cambridge, England

(Received January 8, 1952)

The logical deduction of the existence of the concept of temperature from the laws of thermal equilibrium is examined. The mathematical proof that the relation given by Carathéodory, in his phenomenological discussion of the laws of thermodynamics, is a necessary and sufficient condition is supplied, thus refuting criticism that has been made of his work. The descriptive account sometimes given from a consideration of corresponding isotherms is also examined. It is shown how this qualitative idea can be put into analytical form.

INTRODUCTION

IT is well known that the law of thermal equilibrium between physical systems leads to the concept of temperature. As this is one of the primary concepts of thermodynamics, it is desirable that, in any logical discussion of the laws that can be deduced from the fundamental experiments, this concept be strictly established. The existence of thermal equilibrium between two systems is expressed by an equation of constraint which must be satisfied by the macroscopic variables in terms of which the states of the system are specified. In his phenomenological derivation of the laws of thermodynamics, Carathéodory¹ stated the condition that the equations of constraint satisfy the law of thermal equilibrium. Neither Carathéodory nor more recent expositors² of his viewpoint bothered to give the mathematical proof that this condition is necessary and sufficient—probably because the demonstration is almost trivial mathematically. This has, however, led to quite unwarranted criticism of Carathéodory's work by Zemansky and his colleagues³ who state that "it is doubtful if a general proof can be given without making further assumptions." As these doubts have been expressed, it seems desirable to give the mathematical argument. This demonstration is given in the earlier part of this paper.

By considering the corresponding isotherms of different physical systems, Zemansky and his colleagues also gave an instructive qualitative account of how the concept of temperature

arises, but they did not supplement this by any analytical discussion. Such an analysis is given in the latter part of this paper.

EQUATIONS OF CONSTRAINT

The state of a physical system can be described by a number of macroscopic variables, which can be determined by experiments carried out on it. The experiments of Boyle, Charles, Dalton, Gay-Lussac, and Mariotte on gases showed that the state of a given mass of a particular gas was completely specified by its volume and pressure. When a system is placed, for example, in an electromagnetic field additional variables are necessary—magnetic field and magnetization, electric field and polarization—and likewise, if surface or other effects are of importance, other variables will be needed. The important point is that *at least* two variables are required. The arguments which are given below are quite general, but it seems desirable to state them first in terms of the thermal equilibrium of gases. For, by making the argument specific to a particular physical system in this way, it may be better appreciated by those to whom the meaning of general mathematical expressions is obscure.

The behavior of a physical system depends on the proximity of other systems and the walls which separate them. In the present connection we are concerned with diathermic walls. When two physical systems are separated by a diathermic wall, the values of the parameters which describe their states are not all independent. For example, when two gases are separated by a diathermic wall, and the environment of one of them is changed so that its pressure and volume alter in an arbitrary way,

¹ C. Carathéodory, *Math. Ann.* **67**, 355 (1909).

² M. Born, *Physik Z.* **20**, 288 (1921); A. Landé, *Handbuch der Physik* (Julius Springer, Berlin, 1926), Band IX, Kap. 4.

³ Balamuth, Wolfe, and Zemansky, *Am. J. Phys.* **9**, 199 (1941).

those of the other gas change spontaneously. This implies that the values of the pressures and volumes which specify the states of the two gases are not all independent; that is, there is a functional relationship satisfied by them. This can be written

$$T_1(p_\alpha, V_\alpha; p_\beta, V_\beta) = 0, \quad (1)$$

where the suffixes α and β are used to denote the two gases and T_1 is a function that can be determined by measurements on the system of two gases separated by a diathermic wall.

The spontaneous change continues until the two gases reach an equilibrium state of the combined system. They are then said to be in thermal equilibrium. The equation of constraint (1) is the condition that the two gases are in thermal equilibrium. It is to be noted that the establishment of any kind of equilibrium—mechanical, thermal, chemical, membrane—requires the satisfaction of an equation of constraint.⁴

The equation of constraint is a physically unique relation between the thermodynamical variables of the two gases. It can be written, however, in mathematically different, but equivalent, forms.

TEMPERATURE

It is observed experimentally that if each of two physical systems is in thermal equilibrium with a third system, then they are in thermal equilibrium with one another. This empirical law of thermal equilibrium is the physical basis of the concept of temperature.

Consider three gases A , B , and C and distinguish the pressures and volumes of the different gases by subscripts α , β , and γ . It has been seen that the condition that A and B are in thermal equilibrium is that an equation of constraint,

$$T_1(p_\alpha, V_\alpha; p_\beta, V_\beta) = 0, \quad (1)$$

is satisfied. In order that B and C may be in thermal equilibrium then a second equation of constraint,

$$T_2(p_\beta, V_\beta; p_\gamma, V_\gamma) = 0, \quad (2)$$

must be satisfied. The law of thermal equilibrium

⁴ We have stated these facts very briefly as detailed discussion of these introductory points can be found in the papers already referred to above.

then states that since each of C and A is in thermal equilibrium with B , they are in thermal equilibrium with one another. This requires that there is an equation of constraint,

$$T_3(p_\gamma, V_\gamma; p_\alpha, V_\alpha) = 0, \quad (3)$$

to be satisfied by the thermodynamical parameters of A and C . The law of thermal equilibrium states that the truth of any two of Eqs. (1), (2), and (3) implies the truth also of the third of them. From this fact, we arrive at the concept of temperature by showing that its satisfaction implies that A , B , and C have a property such that its value, measured in terms of the thermodynamical variables of the different gases, assumes the same value when they are in thermal equilibrium.

First consider the mathematical consequences of equations like Eqs. (1) and (2). Equation (1) can be solved for V_β as a function of p_α , V_α , and p_β . This solution can be written

$$V_\beta = \phi_1(p_\alpha, V_\alpha, p_\beta). \quad (4)$$

Equation (2) can be solved for V_β as a function of p_γ , V_γ , and p_β . This solution can be written

$$V_\beta = \phi_2(p_\gamma, V_\gamma, p_\beta). \quad (5)$$

Mathematically, Eqs. (1) and (2) lead to a relation between the *five* variables p_α , V_α , p_γ , V_γ , and p_β .

However, the law of thermal equilibrium requires that the functional relations T_1 and T_2 shall not be indiscriminately arbitrary functions of the variables. It asserts that Eqs. (1) and (2) lead to a unique relation which involves only the *four* variables p_α , V_α , p_γ , and V_γ . In fact, it must be possible to eliminate p_β and V_β together from Eqs. (1) and (2) to yield Eq. (3). It follows that there must exist a function ψ which is such that when the values of $\psi(V_\beta)$ determined from Eqs. (4) and (5) are equated, the terms containing p_β vanish. In this way we get

$$\psi\{\phi_1(p_\alpha, V_\alpha, p_\beta)\} - \psi\{\phi_2(p_\gamma, V_\gamma, p_\beta)\} = 0,$$

which must be equivalent to Eq. (3). This result implies that $\psi(V_\beta)$ determined from Eqs. (4) and (5) can be written

$$\psi(\phi_1) \equiv f_\alpha(p_\alpha, V_\alpha)\psi_1(p_\beta) + \psi_2(p_\beta), \quad (6)$$

$$\psi(\phi_2) \equiv f_\gamma(p_\gamma, V_\gamma)\psi_1(p_\beta) + \psi_2(p_\beta), \quad (7)$$

where ψ_1 and ψ_2 are common to the two forms.

A function $f_\beta(p_\beta, V_\beta)$ can then be introduced by writing

$$f_\beta(p_\beta, V_\beta) = \{\psi(V_\beta) - \psi_2(p_\beta)\} / \psi_1(p_\beta). \quad (8)$$

With Eqs. (6) and (7) this then yields

$$f_\alpha(p_\alpha, V_\alpha) = f_\beta(p_\beta, V_\beta) = f_\gamma(p_\gamma, V_\gamma). \quad (9)$$

This relation is thus established as a necessary condition that Eqs. (1) and (2) shall imply Eq. (3). The appeal to the law of thermal equilibrium lies in the assertion that there exists a function ψ with the appropriate properties.

It requires no elaboration to see that Eq. (9) is also a sufficient condition. The first two members of Eq. (9) give

$$f_\alpha(p_\alpha, V_\alpha) - f_\beta(p_\beta, V_\beta) = 0, \quad (10)$$

which must be mathematically equivalent to Eq. (1). The second and third members of Eq. (9) give

$$f_\beta(p_\beta, V_\beta) - f_\gamma(p_\gamma, V_\gamma) = 0, \quad (11)$$

which must be mathematically equivalent to Eq. (2). Equations (10) and (11) yield, by addition, a relation which involves only p_α , V_α , p_γ , and V_γ as is required, and it must be equivalent to Eq. (3). Thus Eq. (9) is also established as a sufficient condition for thermal equilibrium.

The *necessary* and *sufficient* condition that the truth of any two of Eqs. (1), (2), and (3) implies also the truth of the third of them is that these equations can be put in the form (9), each member of which depends on the thermodynamical variables of only one of the gases. Alternatively, this can be described as the necessary and sufficient condition for the validity of the law of thermal equilibrium. Each of the members of Eq. (9) measures a property of one of the systems. This property assumes a common value, ϑ say, for systems in thermal equilibrium and is called the temperature. The functions f are called the temperature functions; they measure the empirical temperature on a scale which is determined by the choice of the physical system and the processes by which the functions f are constructed. The equations

$$f_\alpha(p_\alpha, V_\alpha) = f_\beta(p_\beta, V_\beta) = f_\gamma(p_\gamma, V_\gamma) = \vartheta \quad (12)$$

are the equations of state of the three gases.

The preceding argument can be put quite generally in the following way. Let $\xi_1, \eta_1, \zeta_1, \dots$; $\xi_2, \eta_2, \zeta_2, \dots$; and $\xi_3, \eta_3, \zeta_3, \dots$ be three sets

of variables each referring to a particular physical system which is specified by the subscripts. Suppose that between these variables there are three relations

$$\left. \begin{aligned} T_1(\xi_1, \eta_1, \zeta_1, \dots; \xi_2, \eta_2, \zeta_2, \dots) &= 0, \\ T_2(\xi_2, \eta_2, \zeta_2, \dots; \xi_3, \eta_3, \zeta_3, \dots) &= 0, \\ T_3(\xi_3, \eta_3, \zeta_3, \dots; \xi_1, \eta_1, \zeta_1, \dots) &= 0, \end{aligned} \right\} \quad (13)$$

and which have the property that the truth of any two of them implies also the truth of the third. That is, the set of variables common to any two of these equations can be eliminated between them to yield the third equation.

This implies that if, for instance, ξ_2 is eliminated from the first and second of Eqs. (13) then all the other variables of the set with subscript 2 must also vanish from the resulting relation. Thus, the set of variables with subscript 2 must appear in T_1 and T_2 in the same algebraic combination; call this $f_2(\xi_2, \eta_2, \zeta_2, \dots)$. The first two of Eqs. (13) thus have the form

$$F_1[\xi_1, \eta_1, \zeta_1, \dots; f_2(\xi_2, \eta_2, \zeta_2, \dots)] = 0,$$

and

$$F_2[f_2(\xi_2, \eta_2, \zeta_2, \dots); \xi_3, \eta_3, \zeta_3, \dots] = 0.$$

These can be solved for $f_2(\xi_2, \eta_2, \zeta_2, \dots)$ and immediately yield the result

$$\begin{aligned} f_1(\xi_1, \eta_1, \zeta_1, \dots) &= f_2(\xi_2, \eta_2, \zeta_2, \dots) \\ &= f_3(\xi_3, \eta_3, \zeta_3, \dots) \end{aligned} \quad (14)$$

as a consequence of the law of thermal equilibrium. The argument for the general case then follows the discussion which is given after Eq. (9).

CORRESPONDING ISOTHERMS

Consider two gases separated by a diathermic wall so that they reach thermal equilibrium. If the state of one of the gases is kept fixed, the pressure and the volume of the other can be varied (but not independently). The locus of the points on a (p, V) -diagram which represents the different states of the second gas is called an isotherm. Each state of the second gas represented by a point on the isotherm is in thermal equilibrium with the fixed state of the first gas. The law of thermal equilibrium then asserts that all the states of the second gas represented by points on an isotherm are in thermal equilibrium with one another. With the second gas in one of these states, the pressure and volume of the first gas can be varied (but not inde-

pendently). In this way an isotherm of the first gas can be drawn.

These two isotherms, one for each gas, represent states of the two gases such that every state of the one gas is in thermal equilibrium with every state of the other gas. This common property which characterizes all these states of the two gases, which are represented by points on a pair of corresponding isotherms, is defined as the temperature. A fuller account of this descriptive introduction of the temperature concept can be found in the paper already cited.³

Our purpose now is to indicate how this descriptive account can be put into analytical terms. Consider two physical systems *A* and *B* which are described by thermodynamical parameters ξ, η, ζ, \dots . The isotherms of the two systems can be determined. Let the parametric representation of the family of surfaces in an indicator diagram which are the isotherms of *A* be

$$\vartheta_\alpha = \phi_\alpha(\xi_\alpha, \eta_\alpha, \zeta_\alpha, \dots), \quad (15)$$

and those of *B* be

$$\vartheta_\beta = \phi_\beta(\xi_\beta, \eta_\beta, \zeta_\beta, \dots). \quad (16)$$

In these equations the parameters ϑ_α and ϑ_β assume different values for the different isotherms. Corresponding isotherms of the two systems can be determined in the way described in the two preceding paragraphs. This method of construction insures that there is a one-to-one correspondence between the members of the two families of isotherms.

Let $\vartheta'_\alpha, \vartheta'_\beta; \vartheta''_\alpha, \vartheta''_\beta; \vartheta'''_\alpha, \vartheta'''_\beta; \dots$ be the values of the parameters which specify corresponding isotherms of the two systems. The existence and identification of pairs of corresponding isotherms mean that we have a set of corresponding values of the two parameters. Given a value of ϑ_α , the value of ϑ_β appropriate to the corresponding isotherm of *B* is determined. This insures that there is a functional relationship between ϑ_α and ϑ_β ; and the one-to-one correspondence between the members of the two sets of isotherms implies that in general it is a single-valued function for one of the parameters as a function of the other. Whether or not this relationship can be expressed in terms of elementary functions is irrelevant in the present connection. All that matters is that it exists.

Let this functional relation be expressed by an equation,

$$\Theta_\alpha(\vartheta_\alpha) = \Theta_\beta(\vartheta_\beta). \quad (17)$$

If the functions (15) and (16) are operated on with Θ_α and Θ_β , respectively, we get

$$\Theta_\alpha[\phi_\alpha(\xi_\alpha, \eta_\alpha, \zeta_\alpha, \dots)] = \Theta_\alpha(\vartheta_\alpha) = \vartheta, \quad (18)$$

and

$$\Theta_\beta[\phi_\beta(\xi_\beta, \eta_\beta, \zeta_\beta, \dots)] = \Theta_\beta(\vartheta_\beta) = \vartheta, \quad (19)$$

where ϑ has been written for the common value of the two members of Eq. (17). If f be written for $\Theta(\phi)$ then, these reduce to

$$f_\alpha(\xi_\alpha, \eta_\alpha, \zeta_\alpha, \dots) = \vartheta = f_\beta(\xi_\beta, \eta_\beta, \zeta_\beta, \dots). \quad (20)$$

The remarks made in the closing paragraph of the preceding section of this paper apply without change to Eqs. (20). These are the equations of state of the two systems and ϑ is the empirical temperature.

CONCLUSION

Each of these arguments used to establish the concept of temperature depends on the law of thermal equilibrium. However, it is made use of in different ways in the two arguments. That given in the third section of this paper makes use of it in the form that if thermal equilibrium subsists between two systems *A* and *B* and also between *B* and a third system *C*, then *A* and *C* are also necessarily in thermal equilibrium. This is the primary form of the law, and it is the way in which it is invoked in all practical measurements of temperature. It is concerned with a single state of each of three physical systems.

The argument given in the fourth section of this paper is concerned with a single infinity of states of each of only two physical systems. The particular states of interest are all those states of the two physical systems such that any state of either system is in thermal equilibrium with every other state of both systems. These are, of course, the states which are represented by a pair of corresponding isotherms on an indicator diagram.

Either of these arguments leads to Eq. (9) as the necessary and sufficient condition for the validity of the law of thermal equilibrium. From this relation the existence of the temperature concept follows immediately by analytical argument.

Remarks on "The Concept of Temperature" by A. R. Miller

H. C. WOLFE

Cooper Union, New York, New York

AND

M. W. ZEMANSKY

College of the City of New York, New York, New York

(Received March 10, 1952)

WE attempted in our paper of 1941¹ to give a satisfactory introduction to the concept of temperature. We did not discuss the existence of a "temperature function," and Dr. Miller has shown in the second part of his paper that it requires only a small extension of our discussion to supply this lack. As he points out, one has only to add that the equation of a family of isotherms for a system whose equilibrium states are described by two variables Y and X can in general be written in parametric form as

$$\Psi(Y, X) = t,$$

where t is the scalar number, called the temperature, which we discussed more fully. We feel that this completion of our argument provides a

¹ Balamuth, Wolfe, and Zemansky, *Am. J. Phys.* **9**, 199 (1941).

satisfactory basis for understanding the temperature concept.

In the first part of his paper, Dr. Miller deals with the Carathéodory theorem on the existence of a temperature function. He criticizes our statement that "neither Carathéodory nor any other author has given a proof of this theorem, and it is doubtful whether a general proof can be given without making further assumptions." He then proceeds to supply several mathematical steps in carrying through the argument from the Carathéodory approach. In so doing, he uses some of the "further assumptions" to which we referred, making assertions "on physical grounds" as to the nature of the functional relationships. While he has not been as explicit as one might wish about these assumptions, it appears that they relate to the single-valuedness of the functions and also, perhaps, to continuity.

Audio-Visual Aids Committee of AAPT

Tape Recordings of Papers Presented at the 1952 Summer Meeting

At the meeting of the Council of AAPT in February, 1952, it was decided to experiment with sound recordings of talks given at AAPT meetings. The Visual Aids Committee was accordingly expanded to become the Audio-Visual Aids Committee. Through the efforts of J. G. Winans a recording was made of nearly every address at the combined meeting of AAPT and the Colloquium of College Physicists at Iowa City June 11-14, 1952.

Tape recordings (single track $7\frac{1}{2}$ in. per second) are available for duplication for reports numbered 1, 2, 3, 8, 10, 11, and 13 through 28, inclusive, the numbers corresponding to the contributed papers whose abstracts appeared in this Journal (*Am. J. Phys.* **20**, 464-468, 1952). In addition, tape recordings are available of the following:

Unesco (15 min). IRA M. FREEMAN

Raydist System (30 min). VERNE F. SWAIM

Demonstration Experiment Round Table. R. H. COOK, H. C. JENSEN, M. J. PRYOR, AND PAUL ROOD (15 min each)

Edge Tones and Whistles (30 min). H. K. SCHILLING

Recent Developments in Geophysical Exploration (30 min). H. B. PEACOCK

Employment of Physicists (30 min). MARSH WHITE
Researches in Magnetism (45 min). R. M. BOZORTH
So-Called Elementary Particles. GEORGE E. UHLENBECK (Introduction by M. M. JAUCH and first 25 min on one 30-min reel)

Duplicates of any of the above-listed talks may be obtained by sending either a reel and postage or the cost of a reel and postage to J. G. Winans, Sterling Hall, University of Wisconsin, Madison 6, Wisconsin.

The cost of reel and postage is \$2.00 per 15 minutes for paper and \$3.00 per 15 minutes for plastic.

Audio-Visual Aids Committee

V. E. EATON, *Wesleyan University*
W. C. MICHELS, *Bryn Mawr College*
R. L. PETRY, *University of the South*
R. H. RANDALL, *City College of New York*
E. M. ROGERS, *Princeton University*
R. L. WEBER, *Pennsylvania State College*
J. G. WINANS, *University of Wisconsin*
M. W. ZEMANSKY, *City College of New York*
(Chairman)

An "Original" Experiment in Heat for the First-Year Laboratory

IAN G. BARBOUR AND RALPH O. KERMAN*

Kalamazoo College, Kalamazoo, Michigan

(Received April 14, 1952)

There is need for elementary experiments in which the "research spirit" is present. Such an experiment involving the measurement of low temperatures has been used in both general physics and physical sciences laboratories. Each pair of students is required to devise its own method for measuring the temperature of a dry ice-alcohol mixture, to design and construct its apparatus, to carry out the measurement, and to report back to the whole group at the close of the laboratory period the following week. Stronger student motivation and insight into basic factors involved in experimental procedure have been among the outcomes of this experiment.

INTRODUCTION

THE perennial discussion of laboratory experiments for first-year physics courses shows that dissatisfaction with the "cut-and-dried" routine of most laboratory periods is not uncommon among physics teachers.¹ In the typical "General Physics" laboratory experiment, the student is presented with a detailed set of instructions indicating what he is to do in each step of the procedure, what formulas he is to use, and often what answer he should obtain. Usually the experiment is designed to "illustrate" a principle or law already discussed in the classroom; the apparatus is often constructed in such a way that various pertinent variables may be read off directly on appropriate scales. Such an experiment certainly contributes to the educational objective of acquiring technical skill in the manipulation of apparatus. A second important objective, that of insight into physical principles and their relation to experiment, is achieved in many laboratory periods, particularly if the instructions and questions are written with this goal in mind; considerable thought and understanding may also be demanded by the nature of the experiment. Subsidiary objectives,² such as facility in writing laboratory reports, acquiring the habits of accuracy, perseverance, and so forth, may also be attained in varying degrees. Yet, there are certain other objectives

to which the typical laboratory experiment makes but a small contribution.

In an actual *research* experiment one is presented primarily with only one element, a *problem*—no ready-built apparatus, no detailed instructions as to how to proceed. Often the major portion of the undertaking consists in planning a procedure which will allow the factors in which one is interested to be investigated in terms of experimentally observable variables, constructing the apparatus to accomplish this end, and modifying the equipment as may be found necessary in the course of construction and operation. This, of course, requires not only creativity and originality, but also clear insight into the physical principles involved and their relation to experimental quantities. *Although a general physics student has neither the time nor the background to carry out an original research project, it is our feeling that an element of challenge and of the research spirit may be present if there are included during the year a few experiments in which no instructions are given and the student has to plan his own procedure, design and construct his own apparatus, and in which he does not know "the correct answer" beforehand.* Since "cookbook" experiments have frequently been deplored and the desirability of more creative situations stressed, but few practical suggestions have been described, this paper will limit itself to a detailed discussion of one such "original" experiment.

EXPERIMENT

This problem in the measurement of low temperatures arose originally in connection with the

* With the collaboration of L. E. Strong, Department of Chemistry, Kalamazoo College.

¹ See for example: W. Thomas, *Am. J. Phys.* **12**, 53 (1944); Lefler, *School Sci. and Math.* **47**, 539 (1947); E. Rogers, *Am. J. Phys.* **15**, 80 (1947); R. W. Lewis, *Science Education* **31**, 14 (1947); G. E. Owen, *Am. J. Phys.* **17**, 270 (1949); F. Cruickshank, *Am. J. Phys.* **17**, 15 (1949).

² M. Kostick, *Am. J. Phys.* **8**, 331 (1940); H. Schlesinger, *J. Chem. Educ.* **12**, 524 (1935).

construction of a continuously sensitive "thermal diffusion" cloud chamber by some advanced students. It was presented at that time to the general physics laboratory, and has proved so valuable that it has since been used each year with laboratory sections in both physics and physical sciences courses, after the topic "Heat" has been treated in the classroom. As the experiment is currently used, the instructor presents the laboratory group with a large Dewar flask containing a mixture of dry ice and a solvent (such as alcohol) and indicates that we need to know the temperature of this mixture. Of course, it is immediately found that all available thermometers run off scale, and that mercury freezes in the mixture. Each pair of students is asked to devise its own method and construct any necessary apparatus to measure the temperature of the mixture. After a preliminary period for individual work on the problem, the group as a whole discusses some of the approaches suggested for its solution. No two teams may use the same method, and the sharing of ideas, with questions and perhaps hints (though never direct suggestions) from the instructor, encourages the emergence of new ways of attacking the problem. Some of the possible methods which have been suggested are outlined below. The remainder of the first two-hour laboratory period is devoted to constructing and testing necessary apparatus. An assistant is present to help with glass-blowing techniques, obtaining materials, etc. He has, however, been instructed to ask leading questions whenever possible, rather than to give direct answers to difficulties that arise.

In the laboratory period *the following week* each team begins by determining the temperature of the dry ice alcohol mixture as accurately as possible with the apparatus constructed the previous week. Then for the remainder of the period students report back to the whole class the method which they employed, the underlying assumptions and principles, demonstrate the apparatus, present their data and calculations, and point out the sources of error and how these errors could be reduced. Each method is then discussed by the class; finally the question "Which method would be the best one to use?" evokes evaluation of the experiments and further comparison of their accuracy.

METHODS OF TEMPERATURE MEASUREMENT

To give a prescribed list of methods for measuring the temperature of the dry-ice mixture would defeat the purpose of the experiment, and the list below is given only to show the breadth of possible techniques which have been suggested by students. Each year the class has developed some new methods or variations not tried previously. The first three are perhaps the most obvious, but even these involve some interesting questions in construction and calibration.

1. *Expansion of a liquid* (alcohol thermometer).
2. *Constant-pressure gas bulb.*
3. *Constant-volume gas bulb.*
4. *Heat transfer* (method of mixtures). The heat capacity of an object is first determined with a calorimeter using temperature ranges measurable with an ordinary thermometer. The object is then left to come to the temperature of the dry ice mixture and is finally transferred to a beaker of water in the calorimeter, the temperature change of which is noted.
5. *Heat transfer* (using heat of fusion of water). A variation of method 4 is to transfer the object, whose heat capacity has been determined, from the dry ice to a beaker of ice water and to weigh the amount of ice which forms on the object. From these data the temperature of the dry ice may readily be calculated.
6. *Linear expansion.* The expansion of a metal bar can be measured using known temperatures, and the relationship extrapolated to dry-ice temperature.
7. *Bimetal strip.* This variation of method 6 allows greater precision but requires considerable ingenuity to capitalize on its inherent accuracy.
8. *Freezing-point of solutions.* Tables showing the freezing point of alcohol-water mixtures and other solutions as a function of concentration are available; by experimenting with various concentrations one can find the composition which just freezes in the dry ice.
9. *Heat conduction method.* A well-insulated bar, the ends of which are at different temperatures, will show an almost uniform thermal gradient. Measuring the temperature as a function of position along such a bar will allow extrapolation to determine an unknown temperature.

10. *Density of a liquid.* An accurate hydrometer facilitates investigation of the density of alcohol or other liquids as a function of temperature.

11. *Thermocouple or resistance thermometer.* If there is a pair of students who have some knowledge from high school of thermoelectricity and want to do some independent reading (or if electricity is discussed before heat in the course) these or other methods may be explored.

ATTITUDE OF INSTRUCTOR

The attitude of the instructor and his assistant is important; they should attempt throughout to give encouragement and ask leading questions, rather than to give direct answers to difficulties that arise. For example, consider the construction of an alcohol thermometer, perhaps the most straightforward of the methods. The pair of students which has decided to use this approach comes to the instructor and asks "What size glass tubing should we use?" The instructor replies "Will it make any difference what size you select?" The students then set about to design their instrument with a view to what they want it to accomplish. After blowing the thermometer, a difficulty arises as to how the narrow stem and bulb may be filled with alcohol; after thinking awhile, they perhaps feel baffled and are convinced that they are unable to solve the problem themselves. Probably all that is needed is some encouragement in analyzing the nature of the problem. ("Why can't you get the alcohol to go down the stem?" "Because there's air in the bulb." "Well how could you avoid that difficulty?" "Hm—by getting rid of the air." Before long they have devised an arrangement for introducing the alcohol while the bulb is evacuated.)

It is often possible in designing and carrying out these experiments, to *anticipate sources of error*, and sometimes to reduce them considerably. (The usual post-mortem consideration of sources of error tends to be a rather routine and uninspiring addition to a completed experiment, included to please the instructor.) For instance the use of Dewar flasks to insure that the mixture is really down to the dry ice temperature during measurement, suggests itself. There is thus a continual succession of educational situations in which the instructor's chief function is to stimu-

late thought. This experiment has been a real challenge to the instructor as well as to the student, and has new interest for him each year because students think of new ways of approaching problems that arise. It should also be mentioned that even with small laboratory sections, this experiment is not an easy one to administer and is rather demanding for the instructor. In laboratory sections run by graduate assistants it could certainly be used only after considerable discussion among the assistants about the objectives of the experiment. Laboratory sections small enough to allow individual attention are also essential.

EVALUATION

Motivation and curiosity on the part of students have been markedly higher than in most experiments; it is a noteworthy occurrence in itself to find many first-year students asking if they may come in to work further on some aspect of a laboratory problem during the week, when they are not required to do so. In addition to a change in student attitude, there have resulted more thought and creativity³ than in most laboratory periods; it is remarkable how much ingenuity and resourcefulness may be evoked from a typical class of first-year physics students. The necessity of relating physical principles to experimentally observable quantities is a valuable experience, as is the design and construction of one's own apparatus. Solution of practical difficulties encountered and understanding of mistakes made have real educational value. Decisions must be made at every stage of the project.

The *discussion of the experiments* among the whole group provides excellent *review of principles* as well as new insights into a wide range of heat phenomena, from the gas laws to thermometry and calorimetry; in addition, it allows the class as a whole to share in the experience of each pair in *experimental procedure*—the difficulties and accomplishments in design, construction, and operation of apparatus in each method. Finally a number of basic questions of *interpretation of data* arise, such as: (a) What *fundamental assumptions* is one making in interpreting the results,

³ See G. Stewart, *Am. J. Phys.* **13**, 291 (1945).

and how might the validity of these assumptions be tested? (e.g., linearity of expansion of a solid, or the gas laws; extrapolation into an unknown region may require an adequate theory of the nature of matter and of temperature); (b) How far is *extrapolation* valid? (several groups have decided to use the freezing point of mercury as an intermediate calibration point); (c) What *probable error* should be assigned to one's result on the basis of the apparatus and standards used? (Probable errors vary widely, from such a relatively crude method as heat conduction

(method 9) to a rather accurate one such as density measurement (method 10). Comparison of the various methods and their results can give a clear understanding of the concept of probable error as a quantity distinct from the *percentage error* of a particular result in comparison with an average or "true" value.)

This experiment has always been enthusiastically received by the class, and the instructors have felt that the expenditure of two laboratory periods has been fully justified by what has been accomplished.

Reproductions of Prints, Drawings, and Paintings of Interest in the History of Physics

48. William Hyde Wollaston and the Discovery of the Dark Lines in the Solar Spectrum

E. C. WATSON

California Institute of Technology, Pasadena, California

(Received July 28, 1952)

That the spectrum of sunlight is crossed by dark lines was observed for the first time by WILLIAM HYDE WOLLASTON in 1802. In celebration of the 150th anniversary of this important discovery, WOLLASTON'S announcement is here reproduced.

THIS year marks the 150th anniversary of the discovery of the dark lines in the solar spectrum. In a paper entitled "A Method of examining refractive and dispersive Powers by prismatic Reflection" and published in the *Philosophical Transactions* II 92, 378-379 (1802), WILLIAM HYDE WOLLASTON wrote as follows:

I cannot conclude these observations on dispersion, without remarking that the colours into which a beam of white light is separable by refraction, appear to me to be neither 7, as they usually are seen in the rainbow, nor reducible by any means (that I can find) to 3, as some persons have conceived; but that, by employing a very narrow pencil of light, 4 primary divisions of the prismatic spectrum may be seen, with a degree of distinctness that, I believe, has not been described nor observed before.

If a beam of day-light be admitted into a dark room by a crevice $1/20$ of an inch broad, and received by the eye at the distance of 10 or 12 feet, through a prism of flint-glass, *free from veins*, held near the eye, the beam is seen to be separated into the four following colours only, red, yellowish green, blue and violet in the proportions represented in Fig. 3. [Reproduced here as Fig. 1.]

The line A that bounds the red side of the spectrum is somewhat confused, which seems in part owing to want of power in the eye to converge red light. The line B,

between red and green, in a certain position of the prism, is perfectly distinct; so also are D and E, the two limits of violet. But C, the limit of green and blue, is not so clearly marked as the rest; and there are also, on each side of this limit, other distinct dark lines, *f* and *g*, either of which, in an imperfect experiment, might be mistaken for the boundary of these colours.

The position of the prism in which the colours are most clearly divided, is when the incident light makes about equal angles with two of its sides. I then found that the spaces AB, BC, CD, DE, occupied by them, were nearly as the numbers 16, 23, 36, 25.

Since the proportions of these colours to each other have been supposed by Dr. Blair to vary according to the medium by which they are produced, I have compared with this appearance, the coloured images caused by prismatic vessels containing substances supposed by him to differ most in this respect, such as strong but colourless nitric acid, rectified oil of turpentine, very pale oil of sassafras, and Canada balsam, also nearly colourless. With each of these, I have found the same arrangement of these 4 colours, and, in similar positions of the prisms, as nearly as I could judge, the same proportions of them.

But, when the inclination of any prism is altered so as to increase the dispersion of the colours, the proportions of them to each other are then also changed, so that the spaces AC and CE, instead of being as before 39 and 61, may be found altered as far as 42 and 58.

WILLIAM HYDE WOLLASTON (1766-1828) was a London physician and chemist, Fellow and Secretary of the Royal Society, who is best known as the discoverer of the elements palladium and rhodium, but whose scientific contributions extended into optics, acoustics, electricity, mechanics, mineralogy, crystallography, astronomy, physiology, botany, and even art. An excellent review of his life and work, together with quite full references, will be found in the *Dictionary of National Biography*. The paper in which he announced the discovery of the solar lines is an important one in several other respects: in it he also described the total-reflection method for measuring refractivity which has been developed by PULFRICH and ABBE, the spectrum of the blue light at the base of a candle flame (now called the Swan spectrum), the blackening of silver chloride by the invisible ultraviolet portion of the spectrum (an observation made independently about a year earlier by J. W. RITTER), and the first observation of the spectrum of an electric spark.

WOLLASTON was an unusually gifted experimenter. His method of viewing the solar spectrum is still a good one for the beginning student to use because of its simplicity. Whereas NEWTON formed a real image of the spectrum on a screen WOLLASTON substituted the dioptric apparatus of the eye for the lens and the retina for the screen. Modern spectroscopes use essentially NEWTON's method, but permit the image to be viewed directly without the use of a screen. As ERNST MACH has pointed out in his *Principles of Physical Optics*, "it is remarkable that Newton, with all his careful experimenting, had not noticed the fixed spectral lines; this would have been easily conceivable if he had experimented only with circular apertures, but he not only recommended, but actually used, a small rectangular slit. The only conclusion that can be drawn is that he and his followers did not work with all the requisite conditions accurately fulfilled at the same time, namely, a sufficiently narrow slit, minimum deviation, and insertion of the screen at the correct distance of the image." It may also have been that the prisms at their disposal were not sufficiently homogeneous and that the discovery of the solar lines was one of those that had to await advances in technology.

While WOLLASTON had a real talent for originating important lines of research, he too often failed to appreciate the true significance of his discoveries and fell short of making a distinct forward step. The discovery of the solar lines is a good example, for, as is clear from his paper, he regarded them as merely the natural boundaries or dividing lines of the pure simple colors of the prismatic spectrum, and it was not until twelve years later when JOSEPH FRAUNHOFER (1787-1826) independently rediscovered them and established their significance that solar and stellar chemistry was started as a vastly rewarding new science. Nevertheless, WOLLASTON must be given credit for being the first actually to see what we now call the Fraunhofer lines.

It may be of some interest to point out just which of the Fraunhofer lines WOLLASTON saw. In his "Report on Optics" to the British Associa-

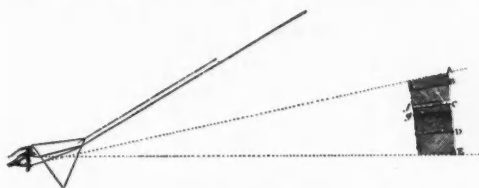


FIG. 1. Diagram of Wollaston's method of viewing the solar spectrum.

tion for the Advancement of Science in 1832, SIR DAVID BREWSTER added the following footnote regarding this matter:

The correspondence of these lines with those of Fraunhofer, I have, with some difficulty, ascertained as follows:

A, B, *f*, C, *g*, D, E, . . . Wollaston's lines
B, D, *b*, F, G, H, . . . Fraunhofer's lines

There is no single line in Fraunhofer's drawing of the spectrum, nor is there any in the real spectrum coincident with the line C of Wollaston; and, indeed, he himself describes it as not being 'so clearly marked as the rest.' I have found, however, that this C line corresponds to a number of lines halfway between *b* and F, which, owing to the absorption of the atmosphere, are particularly visible in the light of the sky near the horizon.

In order to have seen the lines B and H of Fraunhofer, especially the last, Dr. Wollaston's 'beam of day-light' must have come from a part of the sky very near the sun's disc.

Perhaps mention should also be made of the fact that THOMAS YOUNG (1773-1829) immediately repeated WOLLASTON's experiment



FIG. 2. Jackson's portrait of Wollaston.

and used the observations to support his theory that the colors of all natural bodies are similar in their origin to those of thin plates. WOLLASTON's paper was read at a meeting of the Royal Society on June 24, 1802. On July 1, 1802, YOUNG read a paper, also before the Royal Society, at the close of which he stated that he had repeated WOLLASTON's experiments "with perfect success" (*Philosophical Transactions* II,



FIG. 3. Portrait of Wollaston by Sir Thomas Lawrence.

92, 387; also *Lectures on Natural Philosophy*, London, 1807, Vol. II, p. 637). It was only when the lines were viewed under higher magnification and in greater detail, as was done by FRAUNHOFER, that it became obvious that WOLLASTON's interpretation was wrong.

A number of excellent portraits of WOLLASTON exist. Two of these are here reproduced. Figure 2 is from the beautiful mezzotint engraved by WILLIAM WARD, A.R.A., from the painting by J. JACKSON, R.A. Figure 3 reproduces the engraving by F. C. LEWIS made from the portrait sketched by SIR THOMAS LAWRENCE. The originals are magnificent engravings approximately 12 by 15 inches in size.

The [Royal] Institution was also a platform in its early years for lectures other than scientific. The Reverend Sydney Smith gave a very successful set of lectures on morality, and Samuel Taylor Coleridge gave lectures on literature. Gradually the non-scientific lectures gave way to the scientific completely, but as late as 1836 Thomas Carlyle was urged to give a course of lectures on German Literature "at a place called the Royal Institution, where all manner of fashionable people gather in the Spring months to hear such!" But it developed that "the people of the 'Royal Institution' (a kind of sublime Mechanics' Institute for the upper classes) were 'all filled up for this year. . . .'"
—GEORGE A. FOOTE, *Isis*, 43, 7 (1952).

Optical Filters

NORA M. MOHLER

Physics Department, Smith College, Northampton, Massachusetts

AND

JOHN R. LOOFBOUROW*

Biology Department, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received February 28, 1952)

This survey of filters useful in optical work includes the range from ultraviolet through infrared. Descriptions of absorption filters making use of solid, liquid, and gaseous absorbers are treated in Part I. Part II contains descriptions of polarization, interference, and Christian-sen types of filter as well as devices occasionally used in lieu of filters such as focal isolation and residual ray apparatus, gratings, and powder films used effectively as scatterers.

PART I

I. INTRODUCTORY CONSIDERATIONS

1. Uses and Classifications

IN various types of colorimetric, fluorimetric, and other spectroscopic and irradiation problems, it is often necessary to isolate a particular spectral region. The simplest way to remove the unwanted radiation is to use one or more filters. These vary from single pieces of heat-absorbing glass to remove infrared rays or window glass (soda-lime glass) to remove radiation of wavelength less than about 3200A, to the more sophisticated interference and polarization type filters designed to transmit a very narrow spectral region. In recent years a considerable variety of new designs have been described, a number of which are commercially available.

Filters depend upon the following phenomena for their action: (1) selective absorption, (2) scattering of light, (3) interference, (4) polarization, usually in combination with interference, (5) selective refraction and reflection. They may be classified also according to the characteristics of the transmitted beam. A band-pass filter is one that transmits a more or less narrow band of radiation. A "cut-off" or "sharp cut" filter on the other hand transmits either all the radiation of wavelengths longer than those occurring within a certain range (a long-wave-pass filter), or all the radiation of wavelengths shorter than those occurring in a given range (a short-wave-pass filter). Filters that act by selective absorption

may be of any of these types but those that depend on scattering, interference, or polarization are all of the band-pass type.

Filters may be classified also according to the wavelength region for which they are most useful, as infrared, visible, and ultraviolet filters. However, in most cases no sharp dividing line can be drawn. Only if the eye is the detector is it ever safe to ignore transmitted radiation of wavelength longer than 7500 A, or shorter than 4000A, in dealing with the visible region (see Fig. 1). Although in the case of the eye the detector is unaffected by such radiations, they may have to

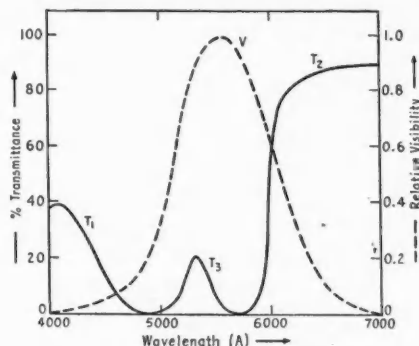


FIG. 1. Transmittance of typical filters for use in the visible region, with the visibility curve of normal photopic vision for comparison. T_1 : Transmittance of Corning glass filter 5113, 3.5–4.5 mm thick. (Thickness controlled to meet color specification 5–58.) T_2 : Transmittance of Wratten gelatine filter No. 25. T_3 : Transmittance of Ilford gelatine filter No. 807 (mercury green). V : Visibility curve (scale on right) from relative luminosity factors for normal light-adapted eye. Data for T_1 , T_2 , T_3 are from manufacturers' literature; see Table V, 11. Data for V are from *Illuminating Engineering Nomenclature and Photometric Standards*, A S A Z 7.1–1942 (Illuminating Eng. Soc., New York, 1942).

* This article was prepared originally as part of a joint project; it has been completely revised since Dr. Loofbourow's death in January, 1951.

be taken into account because of their effects on other components of the apparatus. Either the infrared on the one hand or the ultraviolet on the other may, for example, cause the filter itself to deteriorate.

It should be kept in mind that this division into wavelength regions is an arbitrary one, based on convenience. Experimental methods for the study of different regions vary greatly and as a result one deals observationally with only one region at a time. There is no corresponding theoretical significance to this division. For example, heat-absorbing glass is usually referred to as a cut-off filter of the short-wave-pass type; this nomenclature assumes that the filter is being considered from the point of view of its use in the visible region, for which it is designed. From the standpoint of the three regions—the near ultraviolet, the visible and the near infrared—such types are usually band-pass filters (see Fig. 2). As far as the near ultraviolet alone is concerned the BG 15 illustrated could be described as a cut-off filter of the long-wave-pass type, while for the visible it is a cut-off filter of the short-wave-pass type. For the 3000–8000Å range it is a band-pass filter. In the comparatively limited region of the infrared shown in the figure, there is a single band-pass region, and no inference can be drawn from such data that, if a more extended region were under consideration, there would not be others.

II. ABSORPTION TYPE FILTERS

1. General Consideration of Absorption-Type Filters

Large numbers of substances in the solid, liquid, or gaseous state exhibit selective absorption in various regions of the spectrum. Hence the number of potentially useful materials for absorption filters is quite large. For practical purposes, the number is greatly reduced by criteria of stability and convenience. The absorption characteristics should be stable for a considerable length of time under the conditions of use. A solid filter such as one of glass or gelatine is more convenient to handle than one of a liquid or gas. In general for absorption type filters, as compared to those which depend upon scattering, interference, or polarization effects, the light beam need not be collimated, as the angle of

incidence has little effect on the spectral range or intensity of the transmitted beam. This is particularly advantageous in such applications as the irradiation of substances for the study of photochemical or photobiological effects, since it is usually desirable to obtain the greatest possible energy in such instances.

The possibility of finding a substance with suitable absorbing characteristics is greatest in the visible region of the spectrum. In particular, band-pass absorbing filters of great variety are more readily available in the visible and near ultraviolet regions, down to about 3000Å, than in either the infrared or far ultraviolet. Absorption filters for the infrared usually have extremely broad transmission characteristics so that they isolate a large range of the spectrum rather than

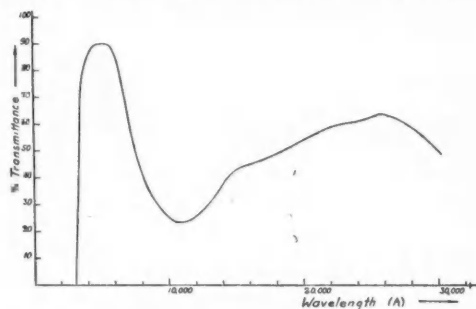


FIG. 2. Transmittance curve of heat-absorbing glass filter, Jena BG 15, 1 mm thick. Data from manufacturer's literature; see Table V, 11.

a narrow region, or else are long-wave-pass filters, eliminating all wavelengths shorter than those of a certain length. In the far ultraviolet region almost all substances otherwise suitable are of the long-wave-pass type.

The stability of the absorption characteristics of a filter is important unless it is to be used for only brief periods, and unless its transmission characteristics can be rechecked frequently. Changes in absorption due to oxidation or polymerization may take place simply on standing, but of more frequent occurrence are changes due to photochemical processes. These become progressively more important as the wavelengths to which the filter is exposed are decreased; they are of little importance in the visible region and of great importance in the short ultraviolet. Many substances otherwise suitable for use as filters

must be ruled out because of rapid decomposition upon exposure to such far ultraviolet radiation. It should also be noted that this radiation is often present even when another spectral region is the one of primary interest.

Absorbed energy may be dissipated by conversion to heat or to fluorescent radiation as well as by photochemical changes. Any one of these processes may be troublesome even though it does not result in changes in the absorption characteristics of the filter. For example, if a filter converts absorbed radiation into heat, the temperature of the filter is raised; this leads to radiation of infrared rays by the filter. Trouble from such radiation is particularly apt to occur if the measurements are made by a thermal type of detector, such as a thermopile. Conversion of absorbed energy to fluorescent radiation occurs primarily in the ultraviolet and occasionally in the visible region. Possible effects of fluorescence radiation on the observations must then be taken into account unless they can be made negligible by some means such as the use of auxiliary filters to absorb the fluorescence-exciting radiation or the resulting fluorescence itself.

Solid filters are more convenient to handle than liquid filters, but the latter may have narrower band-pass characteristics. Liquids or solutions must of course be used in cells that provide appropriate optical path lengths and that are provided with windows transparent to the radiation to be transmitted. Gases have, in general, such low absorption per unit path length that they are of little use, but there are a few particular instances in which they may be employed to advantage.

2. Specifications of Filter Characteristics

When monochromatic radiant energy passes through a homogeneous medium its change in intensity dI is proportional both to its intensity I and the thickness of the absorbing medium dx . That is,

$$dI = -\beta I dx. \quad (1)$$

After integration, if I_0 is the intensity when $x=0$ and I the intensity when $x=d$, the equation may be written in either of the forms

$$I = I_0 e^{-\beta d} \quad \text{or} \quad I = I_0 10^{-Kd}. \quad (2)$$

The constant β is called the coefficient of absorp-

tion and the constant K the extinction coefficient. In either form it may be called Bouguer's law (discovered in 1729) or, more commonly, Lambert's law (announced in 1760). It was noted by Beer in 1852 that changes in concentration of a liquid solution had the same effect as changes in path length, so that Kd may be replaced by kcd , in which c is the concentration and k the absorptivity, or the absorbance per unit concentration per unit path length. In the form $I = I_0 10^{-kcd}$ it may be referred to as Beer's law or as Bunsen's law (announced by Bunsen and Roscoe in 1857).

The validity of Beer's law for mixtures has been queried, but¹ variations seem to be due to disintegration of molecular aggregates, reaction with the solvent, and other modifications of the absorbing material with its concentration. In some cases polarization effects may change the results, however.² A confusing and inconsistent terminology has evolved.³ The continental custom is to use natural logarithms; the English and American custom is to use Napierian logarithms. Photographers are concerned with total effect, Kd , and call this the density, D . The same term may be called the absorbance, and the absorbance divided by the path length the absorbance index (i.e., K). The concentration may be given in grams per liter (mg per ml), or in moles per liter. "Transmission" is being replaced by "transmittance," at least by many authors.

The terminology used in this review may be summarized, with reference to Fig. 3, as follows: $I = I_0 10^{-Kd}$, in which K is the extinction coefficient and d the thickness in centimeters. Transmittance or transmission T is I/I_0 and percent T is $100 \times T$. The density D is defined by the equation $D = \log_{10} I_0/I$. In the case of liquid solutions the absorptivity k and molar absorptivity ϵ are defined by: $I = I_0 10^{-kcd} = I_0 e^{-\epsilon M d}$, in which c is the concentration in grams per liter and M the concentration in moles per liter. If surface reflection is to be taken into account the basic equation becomes $T = \rho 10^{-kd}$, but the definitions given hold if I_0 and I are the intensities measured within the reflecting boundary. Of course, ρ is less than 1.

¹ A. C. Hardy and F. M. Young, *J. Opt. Soc. Am.* **38**, 854 (1948).

² B. Commoner and D. Lipkin, *Science* **110**, 41 (1949).

³ W. R. Brode, *J. Opt. Soc. Am.* **39**, 1022 (1949).

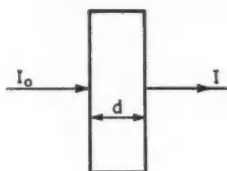
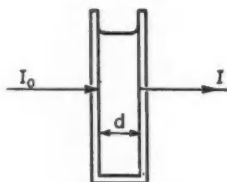


FIG. 3. Diagram to illustrate the terminology used.



The effectiveness of a given filter may be presented graphically as transmittance or as percent transmittance against wavelength, or the necessary data may be given in tabular form. The shape of the transmittance curve changes materially with changes in the thickness of the filter (see Fig. 4). In the case of solid filters, therefore, the filter thickness corresponding to the transmittance curve should be specified.⁴ For liquid

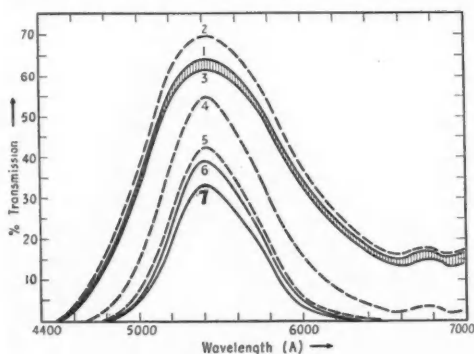


FIG. 4. The effect on the transmittance of a glass filter of the filter thickness, the angle of incidence of the light, and reflection. (1) Transmittance of Corning filter No. 4015, stock thickness (3 ± 0.4 mm). (2) Transmittance of this filter, without reflection. (3) Transmittance assuming light is incident at an angle of 60° . The shaded area represents the variation due to changes in path length for a cone of light with incident angles of 60° – 90° . (4) Transmittance without reflection of material of twice the thickness of (2) and in (5) of three times this thickness. (6) is the transmittance of two filters, each like (1), and (7) is that of three filters. Data for (1) are from the manufacturer's literature (Table V, 11). The other curves are calculated.

⁴ A transmission-thickness nomograph for such calculations is given in *Glass Color Filters* (Corning Glass Works, Corning, New York, 1948).

filters the concentration of the solution as well as the path length must be specified. In gaseous filters the concentration of absorbing molecules depends on the temperature and pressure of the gas so these should be specified as well as the path length, if the meaning of a particular transmission curve is to be clear.

Tabulations of data may be shortened by giving selected information only. A cut-off filter may be defined in terms of the wavelength for which the transmittance is 50 percent or 37 percent (i.e., $1/e$) of the incident energy. For a band-pass filter the wavelength of maximum transmission may be given plus the "half-width" value. By this is meant the wavelength interval between the two points on the transmittance

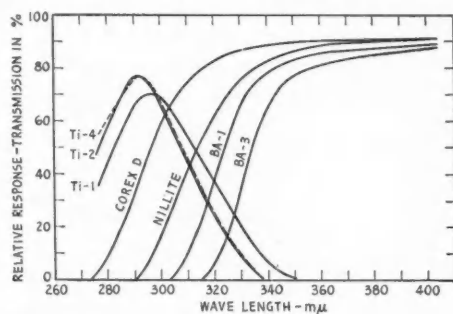


FIG. 5. Response to an equal-energy spectrum of 3 titanium photoelectric cells; spectral transmission of glass filters (Corex D—Corning 9700; Nillite—special window glass; barium-flint glass: Ba-1, 1 mm thick, Ba-3, 3 mm thick). (See reference 8.) From W. W. Coblenz and R. Stair, Research Paper RP 877, J. Research Natl. Bur. Standards 16, 315 (1936).

curve at which the transmittance is half its maximum value. This term has meaning only if the transmittance curve is symmetrical and if the transmittance near the base of the curve is negligible.

In the case of band-pass filters with comparatively narrow but unsymmetrical distribution of transmittance the "spectral centroid" or "effective wavelength" may be given rather than the wavelength of maximum transmittance. If a series of such filters is to be used to transform a colorimeter into a spectrophotometer^{5,6} (Fig.

⁵ D. L. Drabkin, "Spectroscopy: Photometry and Spectrophotometry," *Medical Physics*, ed. O. Glasser (Year Book Publishing Company, Inc., Chicago, Illinois, 1950), Vol. II, p. 1039.

⁶ A. C. Hardy and F. M. Young, *J. Opt. Soc. Am.* 39, 460 (1949).

10), this wavelength rather than the one corresponding to the maximum transmittance is used in plotting curves of the substance being studied. The spectral centroid λ_c is defined as

$$\lambda_c = \frac{\int_0^\infty S_\lambda I_\lambda T_\lambda \lambda d\lambda}{\int_0^\infty S_\lambda I_\lambda T_\lambda d\lambda}, \quad (3)$$

in which, for each wavelength λ , S_λ is the spectral response of the detector, I_λ is the relative spectral energy of the incandescent source, and T_λ is the transmittance of the filter. In practice, integration is replaced by summation over intervals of 50–100 Å.⁷

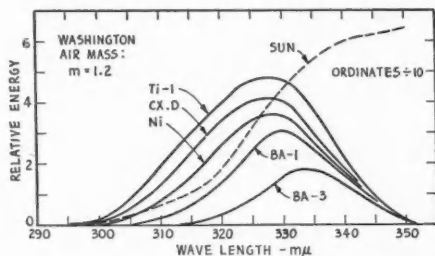


FIG. 6. Spectral-energy curve of the sun as it would be observed with a nonselective radiometer and with the selective photoelectric radiometer Ti-1, without a filter, and with filters Corex D, Nillite, Ba-1, and Ba-3. (See reference 8.)

3. Choice of Filter

The importance of the fact that the response of a given system—source, filter, and detector—depends for each wavelength on the product of the three terms—energy of source, transmittance of filter, sensitivity of detector—cannot be over-emphasized in the problem of choosing a filter for a given purpose. Consideration of the accompanying diagrams concerning apparatus used in the evaluation of ultraviolet solar radiation may serve to make this point clear (Figs. 5 and 6).⁸ The broadness of the resulting energy distributions will also show why, whenever possible, line sources are used with filters rather than incandescent sources. "Negligible transmission" for certain regions for one source may be far from

⁷ K. S. Gibson, *Spectrophotometry* (National Bureau of Standards, Circular 484, 1949), p. 31.

⁸ W. W. Coblenz and R. Stair, J. Research Natl. Bur. Standards 16, 315 (1936).

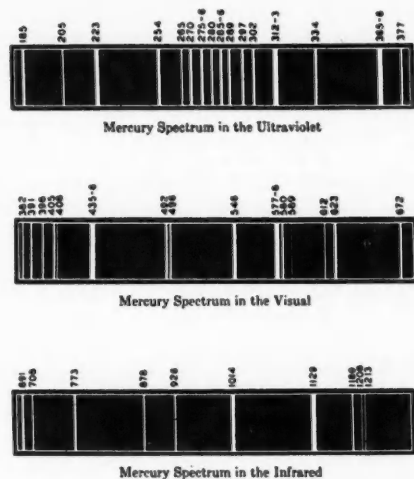


FIG. 7. Schematic representation of the lines of the mercury spectrum. Wavelengths are in mμ; relative intensities are indicated by the width of the lines. Such intensities vary markedly under different arc conditions. From *Wratten Light Filters* (Eastman Kodak Company, Rochester, New York, 1945), p. 21.

negligible with another source or another detector. In Fig. 7 are given schematically the wavelengths and intensities of the spectrum lines of the mercury arc, and in Table I are listed commercially available filters or filter combinations designed for the isolation of lines of such a source.

Region absorbing filters, for transmission of the infrared or the visible or the ultraviolet with absorption of the neighboring regions, are however also of use. Such filters are listed in Table II.

TABLE I. Commercially available filters for the isolation of mercury lines.

Wave-lengths isolated	Manufacturer and code number			
	Corning*	Ilford*	Wratten* (Eastman Kodak Company)	Jena
3650Å	7-83 ^b		18A	WG2+UG2 ^c
4050	5-62			GG13 removes 3650 and shorter
4360	5-74	806 (visual) 805+806+809	50 (visual)	
5460	4-102	807	77	
			77A+58	
5780	3-110	808	22	

* Manufacturers' data (Table V, 11).

^b Color code symbols; their equivalents in currently available filters are: 7-83 is 7380+5860; 5-62 is 3060+4308+5970; 5-74 is 3389+5113; 4-102 is 3484+5120+4303; 3-110 is 3480+4303.

^c E. M. Staats, J. Opt. Soc. Am. 29, 221 (1939).

TABLE II. Commercially available region-absorbing filters.

(1) Transmission of the visible with absorption of the infrared.						
Manufacturer	Code no.	Thickness	Percentage transmittance (maximum)	Wavelength region (Å) of maximum transmittance	Percentage transmittance (cutoff)	Wavelength region (Å) of cutoff
Corning	Aklo Type					
	3961	2.5 mm*	35	5200	0-5	10,000-50,000
	3962	2.5	58	5200	0-20	10,000-50,000
	3965	2.5	74	5200	0-45	10,000-50,000
	3966	2.5	80	5200	0-50	10,000-50,000
Jena	BG-15	2	70	4000-6000	8-45	8000-30,000
	BG-16	2	75	5000	5-30	8000-30,000
	BG-17	3	92	4000-6500	2-10	10,000-30,000
	BG-18	2	90	5000	0-10	6500-17,000
(2) Transmission of the visible with absorption of the ultraviolet.						
Manufacturer	Code no.	Percentage transmittance (maximum)	Wavelength region of maximum transmittance	Percentage transmittance (cutoff)	Wavelength region of cutoff	
Eastman Kodak Co.	Wratten					
	2 (Aesculin)	80-90	4400-7000A	10	3850A	
	4	80-90	5000-7000	10	4600	
Ilford	805 ("Q")	90	4300-7200	5	3800	
Polaroid	UV-375	80-90	4000-7000	5	3500	
	UV-395	80-90	4200-7000	5	3800	
	UV-420	80-90	4400-7000	5	4000	
(3) Transmission of infrared with absorption of visible.						
Manufacturer and code number	Stock thickness (glass filters)	Range of transmittance T (percent)	Wavelength range corresponding to T			
Corning 2540	2.5 mm	0.025-20	750- 9000 mμ			
2550	2	7-73	750-10,000			
Ilford 207		10-50	730- 750			
207+813		1-20	740- 780			
Jena RG 7	2	1-90	750- 1100			
RG 8	2	1-90	680- 710			
Polaroid XRX-60		10-80	800- 900			
XRX-50		10-80	810- 950			
XRX-30		10-80	860- 980			
XRX-20		10-60	880- 1000			
XRX-14		10-45	920- 1040			
(4) Transmission of near ultraviolet with absorption of visible.						
Manufacturer	Code no.	Average thickness (of glass)	Transmittance range (percent)	Corresponding wavelength range		
Corning	5850	4 mm	5-85	4800-3600A ^b		
	9863 (Corex A)	3	5-90	4100-3200 ^a		
Ilford	828 (Chance OX.1)		1-70	4000-3600		
Jena	UG 1	2	5-70	4000-3750		
	UG 2	1	10-90	4000-3500		
	UG 4	2	5-60	4000-3600		
Wratten	18 A		0-80	4000-3800		

^a Stock thickness: 2.15-2.85 mm.^b Transmits also 6900-7500 Å (5-75 percent).^c Transmits also 6800-7200 Å (5-45 percent).

These are often used in connection with other filters of more limiting characteristics.

4. Effect of Reflection and Filter Thickness

Light lost by reflection may be very considerable for a series of filters in tandem. The amount reflected at normal incidence is given by $[(\mu-1)/(\mu+1)]^2$, if μ is the index of refraction of the filter and the filter is in air. For glass of $\mu=1.5$ this amounts to a 4 percent loss at the first surface; for the two surfaces of a single glass filter the transmitted energy is $(0.96)^2$ or about 0.92. Although for most purposes this is negligible, if a series of such filters is used it may lead to excessive loss. Four such filters in tandem, with reflection from eight surfaces, involves a loss in intensity of 28 percent. (See Fig. 4.) The change of percent reflection with change in angle is not appreciable for angles of incidence within about 30° of the normal. For thick filters the change in absorbance due to increased path length is of course more important than it is for thin filters.

5. Effects of Temperature and of Photochemical Action

The effect of temperature changes on absorption filters varies from complete destruction of gelatine filters on heating to near 100°C to negligible shifts in transmittance characteristics of most solid filters when only the day-to-day temperature variations in the laboratory itself are concerned. Changes of the latter type should be taken into account if the filters are used for standardizing purposes, however⁹ (Fig. 8). The shift for large temperature changes may be very considerable. For selenium orange the transmittance curve shifts toward the red; from -180°C to $+430^\circ\text{C}$ the wavelength for 50 percent transmittance changes from about 5400 to 6100 Å. Other glass filters may show changes of comparable magnitude.⁷

Effects on transmittance produced by light, particularly of the far ultraviolet, vary over a wide range. Most filter manufacturers label their products accordingly, so that precautions may be taken if necessary. Of the standards referred

to above (Fig. 8), only selenium orange shows marked change in a short time of exposure.

If the characteristics of the materials used are not known, occasional rechecking should be carried out. A few liquids have proved so desirable in other respects that they are used in constant-flow apparatus. Many filters of gelatine as well as of glass have however been used for years without detectable change. Solarization effects on materials used as "windows" (see Sec. 12) are especially noticeable in the region near 3000Å.¹⁰

6. Absorption Filters for the Visible Region

Two principal types of solid filters are commercially available for use in the visible region: (1) those made of colored glasses, (2) those made by imbibing dyes in gelatine films. For photographic or optical use, gelatine film filters are usually mounted between glass plates to protect them from damage.¹¹ Filters of both of these types can be found with band-pass, long-wave-pass, and short-wave-pass characteristics. The bands transmitted are comparatively wide, several hundred angstroms at half-transmission value. By the combination of suitable filters in tandem it is often possible to obtain band-pass widths of 50 to 200Å, though at the expense

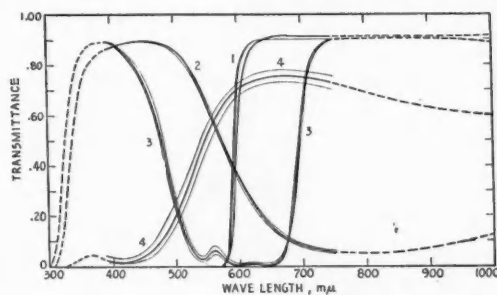


FIG. 8. Spectral transmittance of four types of glass standards issued for the range 390 to 750 $m\mu$: 1, selenium orange (Jena OG 3, 2 mm); 2, copper green (Jena BG 14, 2 mm); 3, cobalt blue (Corning 5551, 3 mm); 4, carbon yellow (Corning 3307, 2 mm). Light lines, standards of different thickness prepared in 1933. Heavy lines, standards in use in 1950. (See reference 9.)

¹⁰ C. Ellis and A. A. Wells, *The Chemical Action of Ultra-violet Rays*, revised by F. F. Heyroth (Reinhold Publishing Corporation, New York, 1941).

¹¹ Wratten filters are protected by a thin coating of lacquer. They can be mounted between sheets of glass and bound with Scotch tape if further protection is needed; they can be obtained cemented in glass.

⁹ K. S. Gibson and M. A. Belknap, J. Research Natl. Bur. Standards 44, 463 (1950).

TABLE III. Commercially available band-pass filters for the visible region.*

Manufacturer	Code number	Maximum transmission (percent)	Wavelength at maximum transmission	Short and long wavelength limits ($T \leq 1$ percent)
Corning	4-64,4010 (4 mm)	25	525 m μ	464, 580
	4-65,4015 (3 mm)	64	545	450 (20 percent at 700)
Ilford	600	7	420	390, 450
	601	29	425	385, 475
	602	14	470	440, 500
	603	12	490	475, 520
	604	10	520	500, 545
	605	5.5	550	530, 585
	606	7.2	580	555, 610
	607	27	600	575, 720
Jena	BG 5 (2 mm)	25	450	(10 percent at 400) 530
	VG 2 (2 mm)	42	525	450, 650
Wratten	36	20	420	(8 percent at 400) 455
	45 (H)	40	480	435, 560
	47 (C5)	51	440	385, 520
	53	19	540	480, 600
	54	1	540	
	61	40	525	480, 610
	74	15	530	510, 550
	75	18	490	460, 520
	B2(58)+H	12	510	480, 530
	G(15)+H	5	525	515, 540

* Data from manufacturer's literature, Table V, 11.

of the maximum percentage of transmission. Such combinations have been designed for isolation of certain spectrum lines, in particular those of the spectrum of a mercury arc, thus making possible the use of this source for monochromatic illumination. Characteristics of a number of useful filters are given in Table III for band-pass and combinations of band-pass filters and in Table IV for cut-off type filters. Many more are tabulated in the *International Critical Tables*, *The Handbook of Chemistry and Physics*, *Physikalisch-Chemische Tabellen*, manufacturer's listings, and other compilations. (See Table V.)

In addition to filters of comparatively standardized properties, gelatine filters for stage lighting are available in great variety. Combinations of these may have very desirable narrow pass-band characteristics. Filters having special properties may be made by dissolving gelatine in a hot solution of an appropriate dye, flowing a film of the gelatine and dye solution onto a glass slide or sheet of Cellophane, and allowing the film to cool and harden. Successive layers may be built up, one on top of the other, in this way and quite large filter areas may be made. Similar filters have been made by soaking No. 600 Du Pont

Cellophane in the desired solution for 12 hours, and then allowing them to dry.^{10,12}

Different lots of the same type of filter may differ considerably in spectral transmission characteristics. In glass filters designed for optical uses an effort is made to compensate for such variation by adjusting the filter thickness of each melt of glass so that the transmission curve corresponds to a standard. If it is important that the characteristics should be known accurately however, published curves should not be relied upon, and the transmission characteristics of the particular filter should be determined. In Fig. 8 are given the measured transmission curves of several identical commercial filters chosen for standard use. Gelatine filters made for decorative purposes are subject to considerable variation, since the amount of imbibed dye in a given area is difficult to control. But those made for optical use are standardized within close tolerances.

Other solid filters for reduction of the intensity of light by constant amounts throughout a considerable range of the spectrum—that is, neutral density filters—are considered later. Those for reduction of light intensities by such proportions

¹² R. Withrow, *Bull. Basic Sci. Research* 3, 82 (1931); U. S. P. 1,938,734 (1933).

TABLE IV.

(1) Typical long-wave-pass filters for the visible and near infrared.				
Manufacturer and code number		λ for which $T=50$ percent	λ for which $T=80$ percent	λ for which $T=0.5$ percent
Corning ^a (glass)	3389 (Noviol A)	430 m μ	511 m μ	405 m μ
	3385 (Noviol C)	492	566	446
	3484 (Noviol D)	545	574	507
	2424	598	629	568
	2404	638	667	608
	2030	670	750	620
	2600	742	800	710
Ilford (gelatin)	608	655	690	615
Jena ^b (glass)	GG 3	430	440	350
	GG 7	475	480	375
	GG 11	500	500	460
	OG 3	580	600	560
	RG 1	600	610	595
	RG 8	700	710	680
Polaroid ^c (plastic)	XXR 60	830	900	760
	XXR 50	860	960	780
Wratten (gelatin)	4 (in glass)	465	500	458
	8 (K2)	495	530	458
	16	535	560	520
	24	595	620	575
	29	625	650	605
	87	790	840	745

Data are taken from the manufacturers' literature. (Table V, 11.) Variations occur in different dye lots, but the characteristic sharpness is maintained.

(2) Solid short-wave-pass filters, visible region.

Manufacturer	Code no.	Wavelength at maximum transmission	Maximum percentage transmission	λ at $T \leq 10$ percent
Jena	UG 3, 2 mm	400 m μ	70	450 (increases to 600)
	BG 1 2	400	92	500 (increases to 650)
	BG 12 2	430	75	480
	BG 18 2	550	72	620
Corning	5970 5 mm	400	30	410 max in uv, 360
	5850 4	400	75	480 max in uv, 360
	5030 5	400	80	520
	5031 4.5	440	82	510

^a Stock thickness 3 ± 1.6 mm. The complete series includes a filter at least every 30 m μ , of similar slope T vs λ ; tolerances from ± 15 to ± 100 m μ .

^b 2-mm thickness.

^c Plastic 0.001 mm thick. XXR 50 is laminated in glass.

in different parts of the spectrum that the radiation from a filament of one temperature may be changed to that characteristic of another temperature—photometric filters—have been developed primarily for photographic and photometric purposes. Data concerning these, as well as concerning filters for use with particular instruments, may be found in the manufacturers' descriptive catalogs.

The number of colored liquids and solutions which have transmission characteristics making them suitable for use as filters in the visible

region is so great that even the most extensive compilations give only partial listings. Many useful solutions may be made from commonly used organic dyes. In the atlases of tables and charts listed in Table V are shown spectral transmissions of solutions and liquids that may be very valuable. A few of the commonly used liquids are listed in Table VI. From such data one may choose filter substances for particular uses. Characteristics of a few solutions that have been found particularly useful in isolating spectral bands are shown in Fig. 9.

TABLE V. Tabulations of filter data.

- (1) *International Critical Tables*, published for the National Research Council (McGraw-Hill Book Company, Inc., New York, 1929); J. Becquerel and J. Rossignol, Vol. V, pp. 268-271, "Spectral Absorption of Light and Heat by Pure Inorganic Substances and Miscellaneous Materials (Nonmetals)"; K. S. Gibson, Vol. V, pp. 271-274, "Spectral Filters"; W. C. Holmes, Vol. VII, pp. 173-209, "The Absorption Spectra of Dyes"; E. P. Carr and M. L. Sherrill, Vol. V, pp. 326-358, "Bibliography of Absorption"; V. Henri, Vol. V, pp. 359-379, "Quantitative Determination of Ultraviolet Absorption Spectra in Solutions of Organic Substances."
- (2) *The Handbook of Chemistry and Physics* (Chemical Rubber Publishing Company, Cleveland, Ohio, 1951, 33rd ed.); "Pigments and Dyes," pp. 2459-2464; "Transmission of Colored Glasses" (includes glasses produced by American Optical, Bausch and Lomb, Chicago Eye Shield, Corning Glass, Jena), pp. 2465-2482; "Transmission of Wratten Filters," pp. 2483-2493; "Transmissibility for Radiations," pp. 2280-2281.
- (3) H. C. Jones, *The Absorption Spectra of Solutions* (Carnegie Institute of Washington, 1905-10).
- (4) Landolt-Bornstein *Physikalische-Chemische Tabellen*, ed. W. A. Roth and K. Scheel (Verlag, Julius Springer, Berlin, 1927). Erster Ergangungsband: "Lichtabsorption im Sichtbaren und im Ultraviolett," pp. 432-462; "Optische Konstanten von Metallen und Legierungen," pp. 463-480; "Optische Konstanten einiger Nichtmetalle," pp. 481-484.
- (5) P. P. Lasareff and P. M. Nedopekin, *Atlas des spectres des substances colorantes* (Academie des Sciences, Leningrad, 1927).
- (6) C. E. K. Mees, *Atlas of Absorption Spectra* (Longmans Green and Company, London, 1909).
- (7) N. Thon, *Annual Tables of Physical Constants and Numerical Data* (National Research Council, Princeton, N. J., 1942).
- (8) F. M. Rowe, Society of Dyers and Colourists, *Colour Index* (Chorley and Pickersgill, Ltd., Leeds, 1924).
- (9) Schultz, *Farbstofftabellen*, ed. L. Lehman (Akad. Verlag, Leipzig, 1930-1940).
- (10) H. S. Uhler and R. W. Wood, *Atlas of Absorption Spectra* (Carnegie Institute of Washington, Washington, D. C., 1907).
- (11) Manufacturer's Data:
- Glass Color Filters* (Corning Glass Works, Corning, New York, 1948),
 - Ilford Colour Filters* (Ilford, Ltd., London),
 - Jena Colored Optical Filter Glasses for Scientific and Technical Purposes* (Jena Glass Works, Schott und Gen., Jena, obtainable from Fish-Schurman Corp., 250 E. 43rd St., New York City),
 - Wratten Light Filters* (Eastman Kodak Company, Rochester, New York, 1945), Revised edition, 1951.
 - Filter Data Sheets* (Polaroid Corporation, Cambridge, Massachusetts, 1950, 1951).

Although such solution filters often isolate a narrower spectral region than solid substances do, they are more apt to be decomposed by ultraviolet light. For this reason a long-wave-pass, cut-off filter may be used in addition, to screen off the photochemically active ultraviolet

TABLE VI. A few useful liquid filters, primarily for the visible region.

Ammonium thiocyanate
Chrome alum (see Fig. 9)
Chrysoidene
Cobalt chloride (in alcohol)
Cobalt sulfate (see Fig. 9)
Copper chloride
Copper nitrate
Copper sulfate (see Fig. 9)
Eosene
Iodine (in carbon tetrachloride)
Lead acetate
Methylene blue
Methyl violet
Neodymium ammonium nitrate
Nickel chloride
Nickel oxide
Nickel sulfate* (see Fig. 9)
Potassium chromate
Potassium dichromate
Quinine sulfate
Sodium nitrite

* For combination cobalt sulfate and nickel sulfate see H. L. J. Bäckström, *Naturwiss.* 21, 251 (1933).

light. For instance in one series of filters to be used in tandem for the isolation of narrow band-pass regions (Fig. 10), Corning No. 395 (3 mm) is placed next to the source, followed by the liquid copper sulfate solution (0.8M CuSO₄, 1 cm) and then by one of the Wratten gelatine filters. (See Table VII.)

7. Absorption Filters for the Infrared Region

The term "infrared" includes the entire range from the limit of the visible region (0.70-0.75 μ) to the microwave region (1 mm). It is customary to refer to the 0.75-2.5 μ radiation as the "very near infrared," to 2.5-25 μ as the "near infrared," and to 25-600 μ as the "far infrared." There are comparatively few substances with good filtering characteristics in any of these regions. This is unfortunate, since in this case it is especially

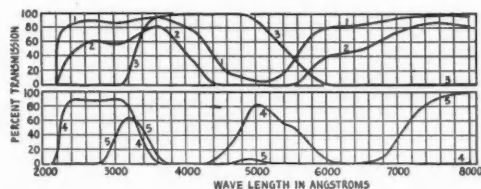


FIG. 9. Light transmission of aqueous solutions of some inorganic salts. CoSO₄·7H₂O, 1-cm path 75 g/l in (1), 300 g/l in (2); CuSO₄·5H₂O, 100 g/l, 5-cm path (3); NiSO₄·6H₂O, 500 g/l (4); and KCr(SO₄)₂·12H₂O, 150 g/l (5); each 1-cm path. From M. Kasha, *J. Opt. Soc. Am.* 38, 929 (1948).

advisable to filter out unwanted radiation, as the thermal detectors usually employed respond to any radiant energy received. This makes the accuracy of experiments peculiarly vulnerable to errors introduced by scattered and reradiated energy. In general, Christiansen and interference filters show more promise than solid filters for the production of narrow bands of radiation (see Part II).

These limitations are less true of the very near infrared region than of the others. Interest during World War II in the transmission of information by infrared radiation with complete elimination of all visible light, resulted in the development of a considerable series of such filters, with a sharp cutoff in the near infrared. Modified Beetle plastic with the Luxol series of Du Pont dyes¹³ and Cellophane, Nylon, or polyvinyl alcohol impregnated with certain green, red, and black

TABLE VII. Characteristics of combination of gelatine filters with Corning No. 395 and CuSO_4 .

Curve*	Wratten filter	Wavelength of maximum (Å)	Wavelength of centroid (Å)
1	29	6200	6250
2	90	5720	5750
3	62	5300	5350
4	75	4800	4850
5	50	4600	4550
6	36	4100	4150

* Numbers refer to curves of Fig. 10.

dyes¹⁴ are characterized by a sharper cutoff than the Corning or Jena glasses made for a similar purpose. They are less subject to injury by overheating than the gelatine Wratten 87. (See Fig. 11. Some of these filters are now commercially available.)

Band-pass filters for use in this region have also been developed.^{14,15} Silver sulfide films on glass have a broad maximum of transmission at about 2.5μ , polymer spray filters and cellulose acetate filters at about 2.0μ . (See Fig. 12.) Both furane-type film and dehydrated high polymers are characterized by a series of somewhat sharper bands. Filters using a conversion of the plastic

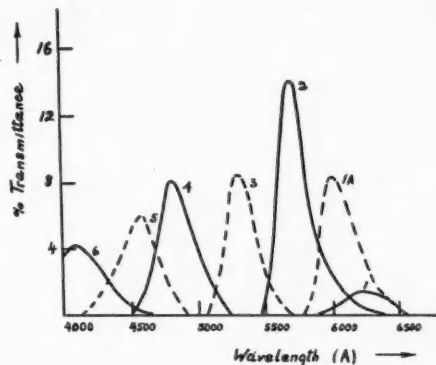


FIG. 10. Narrow band-pass filters for the visible region (see Table VII). Curves redrawn from D. L. Drabkin, "Spectroscopy: Photometry and Spectrophotometry," *Medical Physics*, ed. O. Glasser (Year Book Publishing Company, Inc., Chicago, Illinois, 1950), Vol. II, p. 1039.

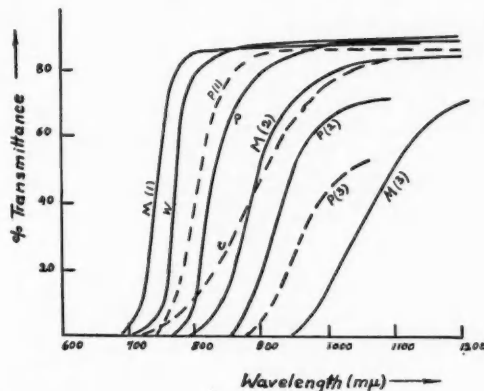


FIG. 11. Very-near infrared long-wave pass filters. *M*: Melmac with (1) Luxol Brilliant Green BL, (2) Pontamine Diazo Blue NA; (3) is (2) with Luxol Fast Green B; *P*: polyvinyl alcohol film with (1) Acid Black G, (2) Rayon Black GS, and (3) Rayon Black; *W*: Wratten No. 87; *C*: Corning No. 2550; *P*: Polaroid XRR 60. From references 13, 14, Table V, 11.

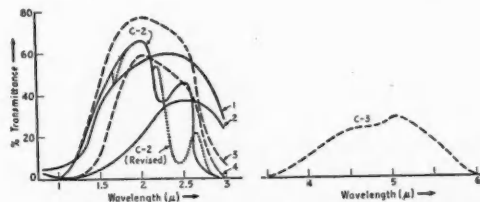


FIG. 12. Very-near infrared band-pass filters. (1) Silver sulfide film on glass with low reflection coating; (2) Silver sulfide film on glass; (3) Polymer spray filter (0.04 mm); (4) Cellulose acetate cast filter (0.13 mm); C-2, C-3 Polaroid C filters. From reference 14 and Table V, 11.

¹³ Shenk, Hodge, Morris, Pickett, and Brode, *J. Opt. Soc. Am.* **36**, 569 (1946).

¹⁴ Blout, Amon, Shepherd, Thomas, West, and Land, *J. Opt. Soc. Am.* **36**, 460 (1946).

¹⁵ Blout, Corley, and Snow, *J. Opt. Soc. Am.* **40**, 415 (1950).

TABLE VIII. Infrared filters.

Substance	Near infrared characteristics 1-30 μ	Far infrared characteristics 30-600 μ
Black paper ^a (.13 mm)	Opaque	Varies with type; Eastman photographic, sharp increase T at 100 μ , 20 percent T 110-150 μ
Camphor soot on celluloid ^{b,c}	Opaque	>75 percent T , 30-130 μ
Celluloid ^c (1 μ)	Banded to 20 μ	>90 percent T , 20-130 μ
Lucite ^a	2-6 percent T , 20-40 μ	>30 percent T , 60-150 μ
Magnesium oxide powder ^{a,d,e}	Opaque to 2 μ ; 40-95 percent T , 3-6 μ	>80 percent T , 40-130 μ
Paraffin ^{a,e} (0.5 mm)	Low transmission, banded to 13 μ	>80 percent T , 50-150 μ
Pliofilm ^a (0.25 mm)	>50 percent, 20-30 μ	Some types banded beyond 80 μ
Polystyrene ^a : Dow (2.6 mm)	Opaque to 20 μ	98-80 percent T , 63-150 μ
U. of Mich. (0.09 mm)	>50 percent T , 20-30 μ	7-30 percent T , 50-150 μ
Potassium bromide	Transmission to 30 μ	>80 percent T , 40-150 μ
Potassium bromide + 5×10^{17} F-centers ^f	Cutoff at 2 μ ; >80 percent T , 3-15 μ	
Quartz ^g (1 mm)	Absorbs 2.5 to 30 μ	Crystalline; >40 percent T , 50-170 μ , banded
Quartz powder ^g (37 μ diam)	Narrow band at 7 μ	Fused >20 percent T , 75-130 μ
Rock salt, unpolished ^a	Filters out 2-14 μ (fine to coarse grind)	Eliminates to 60 μ
Rock salt, powder ^b	Cutoff at 1 μ ; >95 percent T , 3-10 μ	Smaller particles (1-2 μ); 85 percent T , 40-110 μ
Selenium on SiO ₂ ^a	26-80 percent T , 20-30 μ	Max. transmission at 35 μ and 85-110 μ
Selenium powder on nitrocellulose ^{a,g,h}	Cutoff at about 1.5 μ	>90 percent T , 50-115 μ ; 87 percent T , 150 μ
Selenium on pliofilm ^a	40-75 percent T , 20-30 μ	85 percent T , 50 μ ; <20 percent T , 80-130 μ
Silver chloride with stibnite layer	>50 percent T , 1-13 μ	>65 percent T , 40-150 μ
Zinc black on pyroxalin ^a	8 percent T at 20 μ	>40 percent T , 50-150 μ

^a See reference 18.^b See reference 19.^c See reference 21.^d See reference 24.^e See reference 26.^f See reference 23.^g See reference 27.^h See reference 25.

medium into a light-absorbing structure rather than acting by addition of a dye have been prepared from polyvinyl chloride.¹⁶ These absorb the visible and transmit from 20 to 70 percent in the 1 μ to 5.5 μ region, except for a narrow absorption band at 3.3 μ .

Of the Corning infrared transmitting filters (2540, 2550, 2600) only one has a narrow pass-band. Number 2600 transmits about 80 percent at 0.76 μ with cutoffs at 0.710 and 1.1 μ . Similarly Corning filters developed for use in the visible and several Jena filters show distinct transmission bands in the infrared—BG 12, for example, transmits from 0.18 to 0.29 μ with a maximum at 0.26 μ . However the cut-off characteristics are not sharp. Polarizing filters for this region have also appeared. "Infrared Linear Polarizer Type HR" transmits from 0.7 to 2.8 μ , with a maximum at 2 μ .¹⁷

¹⁶ See *Infrared Band-pass Filters, Type C, Data Sheet* (Polaroid Corporation, Cambridge, Massachusetts, 1951).

¹⁷ *Optical Filter Data Sheet* (Polaroid Corporation, Cambridge, Massachusetts, 1950).

Many of the substances that have been reported as useful for the transmission of certain infrared regions with absorption of others vary from sample to sample, so that any tabulation can be only approximate. Films such as camphor soot,^{18,19} gold black,²⁰ and antimony black^{18,21} depend on thickness and method of deposition as well as on material. Typical possibilities are given in Table VIII; reference should be made to the literature for further details. Black paper may be used to absorb both the visible and the near infrared, with excellent transmission in the far infrared. Various types of paraffin have similar characteristics, and Pliofilm, polystyrene, Lucite, and various cellulose films may also be useful.^{18,21}

Use may be made of the difference in amounts

¹⁸ H. S. Seifert and H. M. Randall, *Rev. Sci. Instr.* **11**, 365 (1940).

¹⁹ R. B. Barnes and L. G. Bonner, *J. Opt. Soc. Am.* **26**, 428 (1936).

²⁰ Harris, McGinnies, and Siegel, *J. Opt. Soc. Am.* **38**, 582 (1948); L. Harris, and J. K. Beasley, *J. Opt. Soc. Am.* **42**, 134 (1952).

²¹ R. B. Barnes, *Phys. Rev.* **39**, 562 (1932).

of transmission at different wavelengths, by a drastic change in thickness of material. Quartz, for example, as a thin window is considered transparent to 2.5μ and from 50 to beyond 500μ . If from 3–4 cm thick however, it absorbs the 4– 100μ region, and so acts as a filter to transmit the far infrared. Other materials may be used as short-wave-pass filters: lithium fluoride to 4.5μ , spinel to 5.3μ , corundum to 5.7μ , rock salt to 20μ , potassium bromide to 22μ . (See Table XI.) Many of these have transmission bands in the very far infrared, however. Potassium bromide, sodium chloride, and paraffin have high transmittance in the 400– 600μ region.²²

One device suggested to transform potassium bromide or potassium chloride into effective long-wave-pass filters for the near infrared is to prepare them so they have a stoichiometric excess of the alkali metal. KBr with about 5×10^{17} color centers per cc has a deep visible color, and 50 percent transmission at 2μ .²³

TABLE IX.

(a) Long-wave-pass filters for the ultraviolet.*			
Cellophane filters. (Solvent 1 l water + 125 cc glycerin.)			
W 1	sodium benzoate, 80 g		
W 2	potassium acid phthalate, 40 g		
W 3	cinnamic acid, 30 gr + potassium hydroxide, 10 g.		
Liquid filters.			
K 1	carbon tetrachloride (0.5 cm)		
K 2	potassium acid phthalate 5.0 g/l water, 1 cm		
K 3	naphthalene in isoöctane, 12.8 g/l, 1 cm		
K 4	2,7-dimethyl-3,6-diazacyclohepta-1,6-diene iodide, 20 mg/100 ml water, 1 cm.		
Plastic filters, ^b Polaroid Corporation.			
Designation for 50 percent T			
UV 300	300 m μ		
UV 320	320		
UV 375	375		
UV 395	395		
UV 420	420		
(b) Kasha band-pass filters for the ultraviolet.			
Symbols Fig. 15	Liquid filter (water solvent)	Corning filter	Element as in IXa
A	CuSO ₄ ·5H ₂ O, 100 g/l, 5 cm	5860, 5 mm	K 4
B	NiSO ₄ ·6H ₂ O, 100 g/l, 5 cm	5970, 4 mm	K 3
C ¹	NiSO ₄ ·6H ₂ O, 200 g/l, 5 cm	9863, 3 mm	(c)
C	C ²		K 2
D	NiSO ₄ ·6H ₂ O, 300 g/l, 5 cm	9863, 3 mm	K 4 + K 1
E	NiSO ₄ ·6H ₂ O, 240 g/l, + CoSO ₄ ·7H ₂ O, 45 g/l, 5 cm ^d		

* Symbols as in Fig. 13.

^b Obtainable as film, liquid, or except UV 300 in glass mount.

^c K₂CrO₄, 0.20 g/l, 1 cm.

^d Second element: 1,4-diphenylbutadiene, 4.24 mg/100 ml ether, 1 cm.

²² T. K. McCubbin, Jr., and W. M. Sinton, *J. Opt. Soc. Am.* **40**, 537 (1950); T. K. McCubbin, ONR Report, 1951.
²³ H. Friedman and C. P. Glover, *J. Opt. Soc. Am.* **39**, 795 (1949).

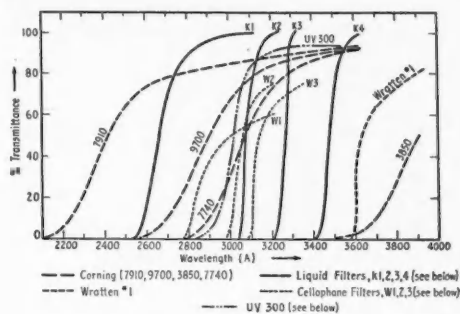


Fig. 13. Long-wave-pass filters of Cellophane, gelatine, glass and combinations of liquids (see Table IX).

The variation of scattering with wavelength may be made use of by preparing the material in powder form. The radiation scattered out of the beam then depends on the size of the powder particles. By making use of the absorption by various thicknesses of the powder and of the sheet which supports it, a variety of filters can be designed. Selenium and magnesium oxide powders on various bases^{24, 25} and quartz powder²⁶ have quite sharp cutoffs. Rock salt may have the surface roughened by grinding or etching, to serve the same purpose.²⁷ The treated polyvinyl chloride previously described with magnesium or zinc oxide fumed on the surface becomes a pass-band filter for the near infrared with a maximum depending on the particle size.¹⁵ Transmission at 5μ is increased from 5 to 30 percent by decreasing the particle size from 1.0 to 0.15μ , for example. Other powder filters will be described in a later section.

8. Absorption Filters for the Ultraviolet Region

As with the infrared, the part of the ultraviolet that lies nearest to the visible region of the spectrum, is richest in the variety of materials with the type of absorption that makes them useful as filters, or as filter components. The absorption due to glass, gelatine, quartz, and Cellophane as well as the chemical activity initiated by far ultraviolet radiation, set a limit of about 2200Å for these devices. It should be noted that quartz and high silica glass become

²⁴ A. H. Pfund, *Phys. Rev.* **36**, 71 (1930).

²⁵ A. H. Pfund, *J. Opt. Soc. Am.* **23**, 375 (1933).

²⁶ J. H. Plummer, *J. Opt. Soc. Am.* **26**, 434 (1936).

²⁷ J. J. McGovern and R. A. Friedel, *J. Opt. Soc. Am.* **37**, 660 (1947).

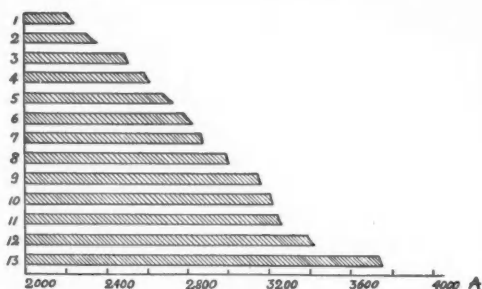


FIG. 14. Absorption of commercial solvents of 1-cm thickness. The slope of the cutoff indicates the sharpness of the absorption. 1 ethyl alcohol, 2 glycerol, 3 acetic acid, 4 ethyl formate, 5 carbon tetrachloride, 6 benzene, 7 xylene, 8 ethyl benzoate, 9 ethyl oxalate, 10 acetone, 11 pyridine, 12 benzyl alcohol, 13 carbon disulfide. Selected and redrawn from W. R. Brode, *J. Phys. Chem.* **30**, 56 (1926).

slightly colored and lose some of their transparency when exposed to very intense short-wave radiation.²⁸

As can be seen from information in Fig. 13 and Table IXa, the solids available for dyeing or for the windows of liquid cells, are themselves long-wave-pass filters. To transform them into broad-band-pass filters, dyes that absorb the visible and transmit the ultraviolet are used. A considerable variety of filters for such use are commercially available. (See Table II, 3.) The number of liquids and solutions that exhibit selective absorption in this region is very large, and extensive compilations exist. (See Table V.) Most of them serve as long-wave-pass filters, however, and many are not stable. A group of liquids that act as long-wave-pass filters or as solvents to produce band-pass filters is shown in Fig. 14.²⁹

Various combinations of cobalt sulfate, copper sulfate, nickel sulfate, and chrome alum have been used for rather broad transmission or in combination for narrower regions. (See Fig. 9.) To isolate a narrow region, filters in tandem must be used. One set recently described³⁰ (Fig. 15 and Table IXb) provides a valuable series of pass-band filters of about 200 Å half-band width. For each one several units are used, and the order is rather important. A 5-cm cell containing

a stable inorganic salt is placed nearest the source; this serves to eliminate the infrared as well as having other filtering uses. This is followed by a Corning filter that eliminates all of the visible as well as part of the ultraviolet. Last comes a 1-cm organic solution, in the position of most protection from the radiations—infrared or ultraviolet—that tend to make it decompose. These filters were evolved primarily for use with a mercury arc.

Metallic films, especially of silver or the alkali metals, have been used as band-pass filters in the ultraviolet. Silver films may be deposited on fused quartz by evaporation techniques to controlled thicknesses. The maximum transmission is at about 3200Å; the percentage transmission of this maximum depends on the thickness of the layer, a variation of 40 to 10 percent with thickness change from 0.02 to 0.055 μ .⁹ Exposed lantern slides, developed, fixed, stripped, and remounted on "Uviol," "Vitaglass," or fused quartz produce filters with a transmission maximum at about 3000Å.³¹ Potassium films on quartz described by Wood³² as thick enough to be opaque in the visible region transmit between 2000 and 3500Å, almost completely at 2500Å. These films are unstable at room temperature. However films can be prepared between quartz-glass plates that are permanent even at 100°C.³³

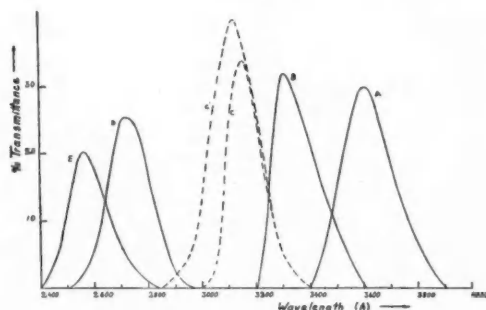


FIG. 15. Band-pass filters for the ultraviolet. Curves redrawn from M. Kasha, *J. Opt. Soc. Am.* **38**, 929 (1948).

³¹ G. C. Brock, *Chem. Abstracts* **27**, 5262 (1933).

³² R. W. Wood, *Phys. Rev.* **44**, 353 (1933).

³³ H. M. O'Bryan, *Rev. Sci. Instr.* **6**, 328 (1935).

Note added in proof:—An alkaline solution of potassium chromate has been reported on as a transmittancy standard for the ultraviolet, from 2000 to 5000Å, in *J. Opt. Soc. Am.* **42**, 441 (1952), by G. W. Haupt.

²⁸ R. J. Dwyer and H. W. Leighton, *Rev. Sci. Instr.* **12**, 156 (1941).

²⁹ W. R. Brode, *J. Phys. Chem.* **30**, 56 (1926).

³⁰ M. Kasha, *J. Opt. Soc. Am.* **38**, 929 (1948).

9. Gaseous Absorption Filters

Gaseous filters are of more frequent use in this region than in any other, but even here are of limited and specialized application. Chlorine and bromine vapor have been perhaps most used, because their transparencies in the ultraviolet have maxima near the mercury lines 2305Å and 2537Å. A 2-cm path length of chlorine at 6 atmospheres have been used to absorb the longer wavelength ultraviolet radiation of a mercury arc.¹⁰ A second broad region of transmittance of chlorine has a maximum in the yellow. Bromine absorption for the visible region is at a maximum in the violet, and in the ultraviolet begins at about 2700Å. The two vapors have been used in a single quartz cell, or in quartz cells in tandem. The combination transmits the 2305Å mercury line, with quite complete suppression of the rest of the ultraviolet.

In some Raman spectra work scattered 2537Å radiation has been eliminated by having mercury vapor in the spectrograph itself, or—more completely—by using a 10-cm cell in which the vapor is produced by heating to 150°C a tube containing a drop of mercury. Since the absorption is followed by re-emission in all directions, the amount of radiation "collimated" in the optical path is negligible.

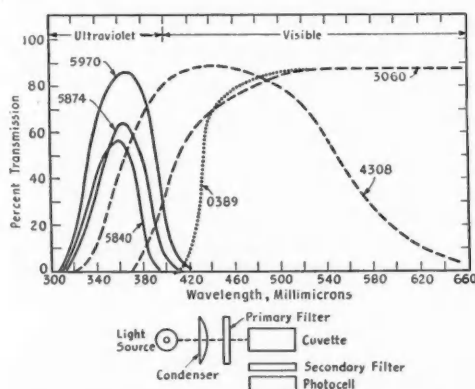


FIG. 16. Filters for fluorescence studies. Transmission curves of filters used in the thiochrome assay of thiamine, with a diagrammatic view of the optical system for fluorophotometry. Filters 5840, 5874, and 5970 are primary filters. 3060, or 0389 followed by 4308, are secondary filters. Adapted from J. R. Loofbourow, "Physical Methods for the Identification and Assay of Vitamins and Hormones," *Vitamins and Hormones* (Academic Press, New York, 1943).

TABLE X. Ultraviolet cutoff as a function of concentration.^a

Compound	Concentration (g/l in isoctane)	Cut-off wavelength
Benzene	1.0	2615 Å ^a
	9.1	2630
	100	2700
	Undiluted	2800
Chlorobenzene	10.0	2750
	100	2810
	Undiluted	2875
Cumene	1.0	2670 ^a
	10	2700
	100	2750
	Undiluted	2920 ^a
Dihydropyran	0.2	2240 ^a
	1.0	2310
	10	2540 ^a
Mesitylene	1.0	2735 ^a
	10	2800
	100	2835
	Undiluted	2900
Phenol	1.0	2800 ^a
	10	2845 ^a
Thiophene ^b	1.0	2480
	10	2540
	100	2590

^a Used with Corning 9863 to remove structure outside the 2350-4000 Å region.

^b Unstable—should be changed every 20-30 minutes.

^c Extrapolated, ignoring minor structure and "tails."

10. Filters for Fluorescence Studies

Several rather specialized uses for ultraviolet filters exist. For fluorescence studies part of the ultraviolet or violet is used. However scattered radiation of the exciting wavelength must be absorbed. Hence a second filter is introduced, chosen for high transmission of the fluorescent light. (See Fig. 16.) Various sets of such filters have been evolved for particular problems.³⁴ It is a very potent technique in the assay of vitamins *B*₁ and *B*₂, for example. In the General Electric Type BH 4 mercury arc, the filter necessary to exclude all but the 3650Å line, is incorporated in the bulb.

In a series of fluorescence studies designed to determine the exciting wavelength region, use was made of the fact that increasing concentration of the solution used as a long-wave-pass filter, increased at a steady and controllable rate

³⁴ J. R. Loofbourow, "Fluorescence," *Medical Physics*, ed. O. Glasser (Year Book Publishing Company, Inc., Chicago, Illinois, 1944), Vol. I, p. 446.

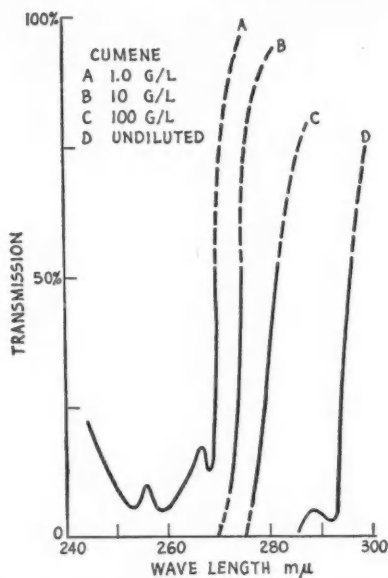


FIG. 17. Cut-off characteristics of cumene and of cumene in isoöctane, as determined by concentration (see reference 35 and Table X).

the shortest waves transmitted. The cut-off wavelengths so determined are given in Table X³⁵ and Fig. 17.

11. Filters for Raman Studies

Perhaps the most stringent of filter requirements is set for their use with Raman spectra. Such long exposures are necessary to show up the Raman lines that even small amounts of radiation not eliminated by the filter are not to be tolerated. A number of combinations have been described for this purpose. A solution thickness of 4 mm of succinic acid circulated in a fused quartz jacket was used to filter out ultraviolet of less than 2300Å wavelength in a study of the Raman spectra of sugars.³⁶ Elimination of the 4047 Å line is possible by coating the pyrex jacket with "dope" (Sample 1809-146-2 of the General Aniline and Film Company)³⁷ with further filtering by a sodium nitrite solution circulated through the filter jacket.³⁸

³⁶ A. M. Bass, *J. Opt. Soc. Am.* **38**, 977 (1948).

³⁷ F. H. Spedding and R. F. Stamm, *J. Chem. Phys.* **10**, 176 (1942), esp. p. 178.

³⁸ G. Glockler and J. F. Haskin, *J. Chem. Phys.* **15**, 759 (1947).

³⁹ C. M. Richards and J. R. Nielsen, *J. Opt. Soc. Am.* **40**, 442 (1950).

A filter that transmits 80 percent of the 4358Å line while passing less than 1 percent of the 4047 Å consists of an alcoholic solution of Rhodamine 5 GDN Extra (E. I. Du Pont de Nemours) circulated through the cylindrical jacket, with a Wratten 2A filter wrapped around the jacket.³⁹ A solution of *p*-nitrotoluene⁴⁰ may be used instead of the 2A filter, the manufacture of which has been discontinued. The Polaroid UV filters should prove particularly useful as the material is obtainable as film or "dope." (See Fig. 13 and Table IXa.) The literature of this subject should be referred to for construction details and for other combinations.

12. Window Materials

If the absorber is in a liquid or gaseous state it is used in an absorption cell of proper length, with plane-parallel end faces. These end windows must of course be as nearly perfectly transparent as possible for the radiation that is to be transmitted and should also be mechanically rigid, chemically stable with respect to both light effects and corrosion by solutions, and should be comparatively easily worked. The same criteria hold for absorbing substances such as plastics or glasses to be impregnated with dye or other absorbing material in the manufacture of solid filters. Window glass can of course be used in the visible region and is transparent to about 3200Å;

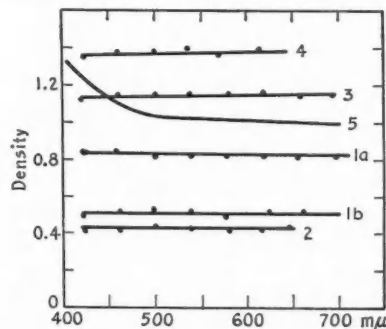


FIG. 18. Neutral density filters. (1) Photographic film, *a* specular and *b* diffuse density; (2) evaporated tungsten; (3) Chromel A evaporated at 10^{-4} mm Hg and (4) at 10^{-4} mm Hg; (5) Wratten filter No. 96. Curves 1-4 from reference 42; 5 from manufacturer's data.

³⁹ B. L. Crawford, Jr., and W. Horwitz, *J. Chem. Phys.* **15**, 268 (1947).

⁴⁰ J. T. Edsall and E. B. Wilson, Jr., *J. Chem. Phys.* **6**, 124 (1938).

TABLE XI. Window materials (thin sections assumed).

Material	Region of high transmission	Comments
Glass	2.5 μ to about 3000 A	Varies with composition
Quartz	2.5 μ to about 1800 A	Both crystalline and fused quartz used
Lithium fluoride (LiF)	50 μ to more than 200 μ 5 μ to 1500 A	In very thin sections used to 7 μ and 1100 A Artificial, to 5 μ
Germanium; ^a silicon	2 μ and beyond	Very pure materials essential
Sodium chloride (NaCl)	to 20 μ	Hygroscopic. For effect of not polishing, see reference h
Potassium bromide (KBr)	to 30 μ	
Fluorite ^b (CaF ₂)	9 μ to 1200 A or less	
Sylvite (KCl)	25 μ to nearly 1800 A	
White sapphire ^c	1435 A into infrared, no bands 1850 A to 2.5 μ	More easily worked than other windows in parts of range
Pliofilm, polystyrene and cellulose nitrate ^{d,e}	20 μ to >40 μ	
Silver chloride (AgCl) ^f	5-20 μ	Darkens if exposed to light <4500 A. Coat with AgS, Se, or stibnite-antimony sulfide
Tellurium ^g	4.5-10 μ	

^a H. B. Briggs, J. Opt. Soc. Am. 42, 686 (1952).^b D. C. Stockbarger, J. Opt. Soc. Am. 39, 731 (1949).^c Freed, McMurry, and Rosenblum, J. Chem. Phys. 7, 853 (1939).^d H. S. Seifert and H. M. Randall, Rev. Sci. Instr. 11, 365 (1940).^e D. C. Stockbarger, Disc. Faraday Soc. No. 5, 299 (1949).^f H. C. Kremers, J. Opt. Soc. Am. 37, 337 (1947); M. Hyman and B. Billings, J. Opt. Soc. Am. 37, 113 (1947); P. Fugassi and D. S. McKinney, Rev. Sci. Instr. 13, 335 (1942).^g J. J. Loferski and P. H. Miller, Jr., Abs. H-1, Bull. Am. Phys. Soc. 26, No. 4 (1951).^h J. J. McGovern and R. A. Friedel, J. Opt. Soc. Am. 37, 660 (1947); see also van Zandt Williams, Rev. Sci. Instr. 19, 135 (1948).

high silica glass can be used considerably farther into the ultraviolet, and quartz, either crystalline or fused, to about 1800A.¹⁰ Some substances are used in the infrared and far ultraviolet in spite of undesirable characteristics because their transparency is unusually good or even unique. Rock salt is commonly used in the infrared, for instance, in spite of its susceptibility to attack by atmospheric moisture. Silver chloride is excellent in its transparency in the 3-14 μ region but has high reflectance and breaks down on exposure to light if wavelength is shorter than 4310 A. A coating with evaporated stibnite prevents the latter, and a low reflection plastic coating reduces the former defect.⁴¹ Characteristics of some of the more usual window materials are given in Table XI.

13. Neutral Filters

It is occasionally of advantage to be able to remove a definite percentage of incident radiation, without changing the relative proportion of

different wavelengths. Photographic film may, for example, be exposed, developed and fixed, to serve as such a filter. Although excellent as far as neutrality is concerned, the scattering of the silver particles makes the actual density a function of the collimation of the light beam. (See Fig. 18.) It is however difficult to get a "true" gray for any extended region of the spectrum. For perhaps the best standards use is made of a more or less fine wire mesh, which is vibrated to avoid diffraction pattern effects. Such screens calibrated for the near ultraviolet give comparable results when used in the vacuum ultraviolet.⁴² Neutral filters of densities 0.1 to 1 in 0.1 steps, and in addition those of densities 2.0, 3.0, and 4.0 are available in the Eastman Kodak Company Wratten Filter series. The curve of Number 96 (about 10 percent transmission) is shown. (Figure 18.) More accurately neutral are the Chromel A filters⁴³ evaporated at 10⁻⁴ mm Hg. These are effective throughout the visible region to densities of 2.

⁴¹ M. Hyman, Jr., and B. H. Billings, J. Opt. Soc. Am. 37, 113 (1947).

⁴² Harrison, Gaddis, and Coffin, J. Chem. Phys. 18, 221 (1950).

⁴³ M. Banning, J. Opt. Soc. Am. 37, 686 (1947).

NOTES AND DISCUSSION

An Electronic Demonstrator for Damped Electric Oscillations

W. G. MARBURGER

Western Michigan College of Education, Kalamazoo, Michigan

WHEN a charged capacitor is discharged through a circuit consisting of an inductor and a resistor, the discharge current is oscillatory, provided $R < 2\sqrt{L/C}$. This oscillating discharge current has a damped wave form, with the damping factor $\alpha = R/2L$. This wave form can be displayed quite well on the screen of a cathode-ray oscilloscope. To do this the capacitor must be pulsed periodically at short but regular intervals.

A motor-driven mechanical commutator can be used to produce the charging pulses. Such a device is difficult to adjust and even more difficult to hold in synchronism with the sweep of the oscilloscope. Furthermore, mechanical commutators often produce troublesome electrical transients which detract from the quality of the damped wave pattern displayed on the oscilloscope screen. An electronic demonstrator for pulsing or "ringing" an L - C - R circuit to show the damped oscillatory character of the discharge current is described in this note.

The schematic circuit diagram for this demonstrator is shown in Fig. 1. A conventional 60-cycle transformer-rectifier-filter circuit supplies about 300 volts dc for the plate circuit and 0 to -2.0 volts for the control-grid bias of a type 2050 thyatron tube. The screen grid of this tube is connected to the control grid. The L - C - R test circuit is inserted in series with the cathode, as shown. Voltage for driving the Y amplifier of the oscilloscope is obtained directly across the capacitor C , between the cathode and ground.

When the voltage applied to the plate of the thyatron builds up to the ignition point, the tube fires and the charge on the capacitor C' passes through the tube to the condenser C . This occurs very rapidly so that the charging current is a sharp pulse of short duration. Because the resistor R has a high value, the voltage of the plate of the thyatron momentarily drops when the tube fires and the arc through the tube is extinguished. The capacitor C then discharges through the inductor L and the resistor R . In the meantime, the capacitor C' charges again through the resistor R' to the ignition point, and the tube is ready to fire and produce the next charging pulse for the capacitor C .

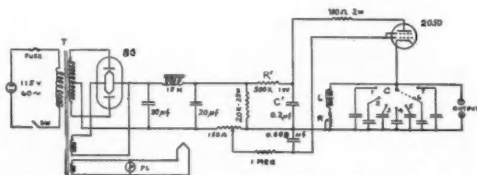


FIG. 1. Schematic circuit diagram of the demonstrator. Values of components not indicated in the diagram are: C_1 , 0.01 μ f; C_2 , 0.02 μ f; C_3 , 0.04 μ f; C_4 , 0.1 μ f; C_5 , 0.2 μ f; C_6 , 0.35 μ f; C_7 , 1.0 μ f; L = 100 mH; R = 400 ohms, variable; T , a Thordarson 22R30 transformer.

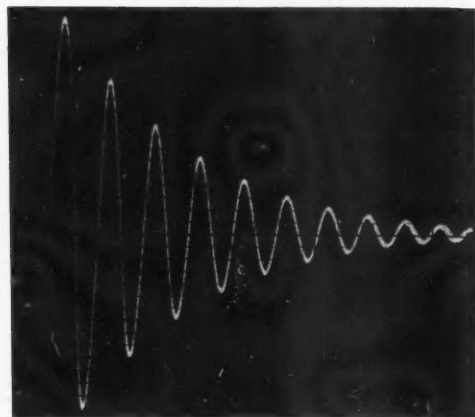


FIG. 2. Photograph of a typical pattern on the oscilloscope screen. An intensifier was used on the negative to improve the contrast.

The time between pulses is determined in part by the time constant $C'R'$. In this demonstrator both C' and R' are fixed in value. The firing voltage is controlled in part by the negative bias of the control grid of the thyatron. This bias voltage can be varied within limits by means of a potentiometer R_2 .

A capacitor-selector switch permits the operator to select any one of 7 different values of capacitance for C . A rheostat control permits a change of the value of R through a range of 400 ohms from a minimum value fixed by the resistance of the inductor. The inductor is a multi-layer air-core coil whose inductance is 100 millihenries. This inductance can be varied by inserting different types of metal rods into the coil.

A Dumont Type 304H oscilloscope is used with this demonstrator. Best results have been obtained with a driven sweep, the initial charging pulse being used to trigger the sweep. With a little care in making adjustments it is possible to obtain a damped wave pattern that will remain practically stationary on the screen for many minutes at a time. Figure 2 shows the type of pattern obtained. A very satisfactory pattern can also be obtained with a linear sweep of the oscilloscope beam, synchronized at 60 cps.

Conversion of an Optical Spectrometer for X-Ray Problems

ROBERT H. McFARLAND

Kansas State College, Manhattan, Kansas

THE need frequently arises in a physics laboratory for a single-crystal x-ray spectrometer capable of measuring x-ray spectra using counting techniques. The construction of such an instrument is laborious because

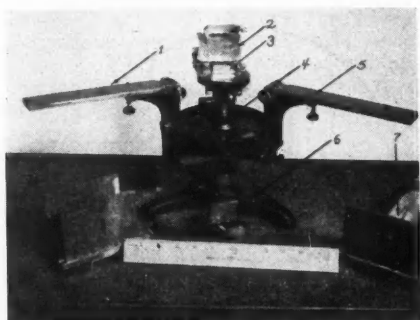


FIG. 1. A converted optical spectrometer which was used for x-ray measurements in the energy region of twenty to one hundred kilovolts. The working parts are indicated by number as follows: 1, 5, Ways for holding slit systems. 2, Crystal and crystal holder assembly. 3, Mirror for optical lever measurements. 4, Slits *B* and *C* together with slit holder. These were combined and moved as a single unit. 6, Lead wedge placed on the crystal table near the surface of the crystal to reduce spurious scattering. 7, Assembly consisting of *G-M* tube, holder, and slit *D*. This is fastened to way 5.

of the degree of accuracy needed in the preparation of bearing surfaces. An instrument adequate for many purposes may be prepared from an optical spectrometer by removing the optical system and by replacing it with simple ways for supporting the collimating slits and the Geiger-Mueller tubes.

In use, the angle through which either the crystal or the *GM* counter moves may be measured by use of the spectrometer's goniometer. The remaining angle is measured by means of an optical-lever system with the movable mirror attached to the appropriate portion of the spectrometer. In this conversion, the goniometer was used for measuring the *GM* counter angle. The least count of the goniometer's vernier was one minute. However, it was found possible to estimate this reading to one-half minute; and since other spectrometers are available with least counts of twenty seconds or less, it is possible to extend this accuracy if necessary. By using an optical lever with a three-meter lever arm and reading the position of the index to one-half millimeter, it was possible to measure the crystal angle with an accuracy somewhat greater than one-half minute. Since the mechanical conversion is extremely simple and different for each spectrometer type, no detailed drawings

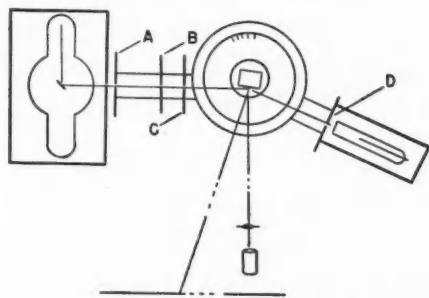


FIG. 2. A schematic diagram of the spectrometer indicating the relative positioning of the slits.

are presented. However, the conversion for a specific type may be seen in Fig. 1.

Figure 2 indicates the relative positions of the component parts of the spectrometer. Defining lead slits at *B* and *D* are one-mil slits. Slits *A* and *C* are 5 and 250 mils wide, respectively. Slits *B* and *D* are 10 inches from the point of rotation of the unpolished calcite crystal. Slit *B* is approximately 8 inches from *A* and 4 inches from *C*. A lead wedge shown in Fig. 1 was incorporated into the crystal table to reduce spurious scattering effects at the crystal surface. Adequate shielding must be provided at the x-ray tube and elsewhere to prevent over-tolerance exposure of persons in the vicinity of the apparatus.

Typical spectra observed with a tungsten target gave an estimated half-width of 0.2 kv for the $K\alpha$ line.

Note on the Behavior of a Certain Symmetrical Top

J. A. JACOBS

University of Toronto, Toronto, Canada

CONSIDERABLE interest has been aroused in the seemingly strange behavior of a small plastic top which, after spinning in a normal manner for a short time, turns upside down and spins on its stem. It is the purpose of this short note to show that this behavior is no more unexpected than many similar phenomena which occur when a body is spun and becomes subject to gyroscopic action. A complete mathematical analysis is lengthy and only the main results, which, of course, are not new, are discussed here.

The top was first produced in Denmark and was later manufactured in Canada, being obtainable at any drug store or tobacconist for 25 cents. The top is illustrated in Fig. 1. It consists of a hollow sphere $1\frac{1}{2}$ in. external diameter with approximately $\frac{1}{2}$ in. sliced off it and a stem just over $\frac{1}{2}$ in. long attached to it. The top weighs a little over $\frac{1}{4}$ oz.

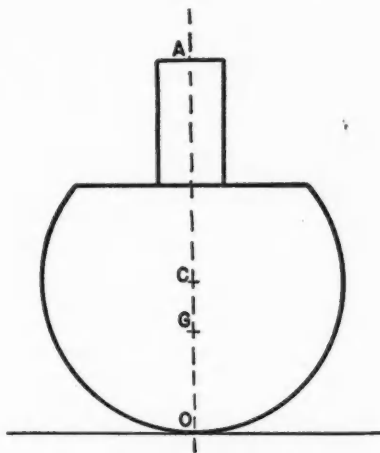


FIG. 1. Hollow symmetrical top.

The first important fact that an examination of the construction of the top reveals, is that, when it is upright, the center of gravity G is below the center of curvature C of a meridian cross section. This means that the top will remain in stable equilibrium with its axis vertical when it is at rest on a horizontal surface (as in Fig. 1), but that when it is spun, motion with the axis vertical is unstable. The converse is of course true for a conventional top. This instability is demonstrated by placing a prolate spheroid with its long axis horizontal on a horizontal plane and giving it a spin about a vertical axis. The spheroid will not continue to spin with its long axis horizontal, but will raise its center of gravity until it is spinning with its long axis vertical. The same phenomenon occurs when a loaded sphere is spun. If the initial spin is sufficiently great and the plane sufficiently rough, the sphere will ultimately spin with the loaded part at its greatest distance from the plane.

Figure 2 shows the top spinning with its axis inclined at an angle to the vertical. Force R is the normal reaction at the point of contact O' with the plane, whilst F_1 and F_2 are the components of the frictional force, the former acting along the intersection of the horizontal plane and the vertical plane CGO' , and the latter acting at right angles to the vertical plane as shown. If the top is spun from the hand onto the plane, then initially the point of contact O' slips back on the plane. The magnitude of the frictional force will depend on R and is not limited to that required for pure rolling. The component F_1 provides the necessary normal acceleration for G to describe a circle, whilst the component F_2 increases the rate of precession and as a consequence causes the center of gravity to rise. For a conventional top the motion is started with the center of gravity above the center of curvature, and since the frictional component F_2 on the rounded apex will cause the center of gravity to rise, the axis rises towards the vertical, approaching the position of sleep. For the special top considered here, however, the first stage of the motion is approximately the reverse of that of a sleeping top. The center of gravity is initially below the center of curvature and the raising of the center of gravity means that the axis moves away from the vertical. After the horizontal position of the axis is reached, the rest of the motion is approximately that of the sleeping top.

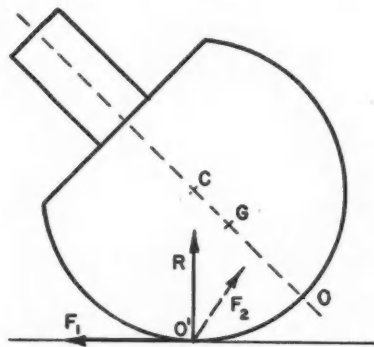


FIG. 2. Top spinning with its axis inclined.

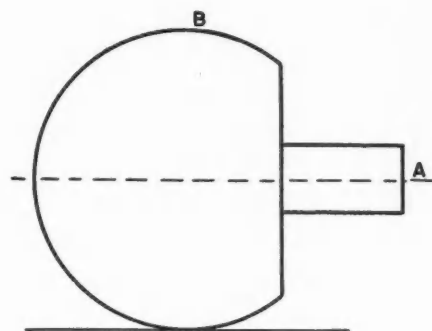


FIG. 3. Top held in horizontal position.

The motion of the top furnishes a good illustration of the wandering of the instantaneous axis in the body. The resultant spin is always in the same sense about a roughly vertical axis, so that when the top is inverted the spin is in the same sense in space, but reversed in the body. It is interesting to note that the motion can be started half-way. If the top is held in the position indicated in Fig. 3 and struck horizontally at A , at the same time holding a finger lightly pressed at B , then the motion is exactly that obtained by spinning in the usual way, after the axis has reached the horizontal position. This shows that in the latter position, there is no spin about the figure axis. The last part of the motion is thus similar to that of the spinning of the prolate spheroid already discussed.

From a consideration of the angular momentum and energy of the top, we can obtain conditions necessary for the rise of the top to the position where G is vertically above C . For the top considered here, the ratio CG/CO is small, and it can be shown that the ratio A/C must lie between rather fine limits, where C and A are the moments of inertia about the axis of the top and about a perpendicular axis through G . It is obvious from the construction of the top that A does not differ appreciably from C . Another condition is of course that the initial spin must be greater than a certain minimum value. The spin is considerably reduced by the time the top has reached the inverted position—its spin when the axis is horizontal being little more than half its initial value. There is a further interesting condition that must also be satisfied, *viz.*, that the initial energy of the top must be reduced by sliding friction to an absolute minimum. Without a sufficient amount of sliding, the top would not rise. It is easy to verify that the prolate spheroid already discussed would not rise if its surface were very rough.

Physics Library Expenditures

WALLACE A. HILTON AND OPAL R. CARLIN*
William Jewell College, Liberty, Missouri

AS a means of determining present physics-library expenditures, a questionnaire was sent to the chairmen of physics departments in 249 selected colleges and uni-

versities. Replies were received from 179 or 72 percent of these institutions. The highest degree offered in 42 of these schools was the Ph.D. In 49 institutions the master's degree was the highest offered in physics while 75 schools granted only a bachelor's degree. Thirteen were not offering a physics major.

Of the 42 institutions offering the Ph.D., 25 were state supported schools and 17 were privately supported institutions. The average number of students in these schools working toward some degree in physics was 98, the range being from 39 to 309. Thirty-four schools reported an average budget for books of \$725, the extremes being \$250 and \$1950. Twenty of these schools reported an average of \$590 spent per year for periodicals, the range being from \$75 to \$1600.

TABLE I. Summary of yearly budgets for physics books and periodicals in the three types of degree-granting institutions.

	Ph.D.	M.S.	A.B.
1. Number of schools reporting	42	49	75
1. State institutions	25	18	5
2. Private institutions	17	24	61
3. Teachers colleges		3	9
4. Municipal institutions		4	
2. Average number working toward degree in physics	98	29	12
1. Range of students	39-309	6-170	0-75
3. Average budget for books	\$725	\$342	\$155
1. Range	\$250-\$1950	\$85-\$1000	\$15-\$1000
4. Average budget for periodicals	\$590	\$193	\$83
1. Range	\$75-\$1600	\$40-\$500	\$15-\$300

Of the 49 schools offering a master's but no higher degree in physics, 18 were state universities, 3 were teachers colleges, 24 were privately supported institutions, and 4 were municipal institutions. Forty-seven schools reported an average of 29 students working toward a degree in physics, the range being from 6 to 170 students. Forty-two departments reported an average budget of \$342 for physics books with a range of from \$85 to \$1000. Thirty-six schools reported an average budget of \$193 for periodicals, the low and high figures being \$40 and \$500.

Seventy-five institutions granting a bachelor's degree in physics responded to the questionnaire. Sixty-one of these were privately supported institutions, 5 were state supported schools, and 9 were teachers colleges. There was an average of 12 students per school working toward a degree in physics, the range being from 0 to 75. The average budget for books was \$155 with the range being from \$15 to \$1000. The average yearly budget for periodicals was \$83 with a range of from \$15 to \$300.

A summary of types of institutions reporting, together with the average yearly budget for books and periodicals is given in Table I.

* Librarian, William Jewell College.

A New Design for a Nodal Slide

LEONARD EISNER*

Pennsylvania State College, State College, Pennsylvania

IN the usual type of nodal slide the image gets out of focus as the lens carriage is shifted; also for mechanical reasons the carriage is short and thus limits the size of the system on test and the location of its nodal points. The design below achieves advantages in these respects by adjusting the position of the axis of rotation rather than the lens system itself. An optical bench, positioned on frame *A* (Fig. 1), is rotated about the center of disk *B*

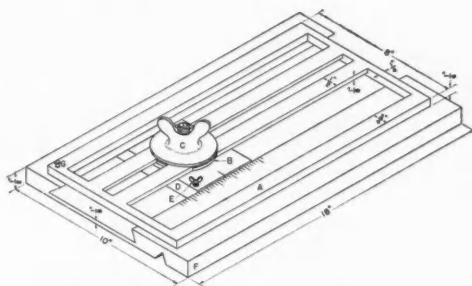


FIG. 1. Base arrangement to adapt optical bench for use as a nodal slide.

which fits snugly in a hole in *D*. The disk *B* has slots to accommodate the central rails of *A*. Thumbnut *C* fits a screw projecting centrally from *B* and clamps the axis at any desired position. The carriage *D* slides in beveled ways and is locked by thumbscrew *E*. The baseplate *F* may be grooved to fit a long optical bench which also supports the source and screen or microscope. Pin *G* locks *A* to *F* when the axis is changed. Axis settings are read on a scale on *F*; however, the fiduciary line on *D* need not be exactly opposite the center of *B* since the *F* scale reading is translated into an accurate value of axis position on the optical bench scale (the correspondence between scales being established by plumb line to the center of *B*, or by measurements on a lens of known characteristics, etc.).

This design is useful for laboratory experiments¹ or demonstration, or for precise measurements, e.g., of aberrations.² One can test a heavy optical instrument, or a telephoto lens with principal planes far in front of the lens; or show the effects of different media by means of a lens at one end of a water trough. In demonstrating a nodal slide it is a good idea to place the source in the focal plane (rather than at infinity as is usual) and to show that the image on a distant screen remains stationary when the axis of rotation is at the first nodal point.³

* Now at Battelle Memorial Institute, Columbus, Ohio.

¹ Taylor, *College Manual of Optics* (Ginn & Company, New York, 1924), p. 9.

² Hardy and Perrin, *The Principles of Optics* (McGraw-Hill Book Company, Inc., New York, 1932), pp. 377-381.

³ This arrangement was suggested to the author by Professor F. W. Sears.

LETTERS TO THE EDITOR

Some Observations on the Electrostatic
Attraction of a Stream of Water

IT is well known that a fine stream of water from a tap can be attracted by an electrostatically charged rod. In demonstrating this to a small class, the question arose as to whether the water had been charged by some frictional effect as it left the tap. While demonstrating that either positive or negatively charged rods produced attraction, an unusually large charge was placed on an ebonite rod and held horizontally near the stream. The stream broke up into fine droplets some of which ended up by following a spiral path with the rod as its axis. The effect is curious to observe and the mathematical analysis of the path taken by the droplets is challenging.

UNESCO Technical Assistance Mission,
University College,
Baghdad, Iraq

ALBERT V. BAEZ*

* On leave from the University of Redlands, Redlands, California.

Concerning the Frequencies Resulting
from Distortion

THIS is in answer to the question raised by Howard N. Maxwell in the May, 1952 issue of the *American Journal of Physics* concerning the physical origin of the harmonic frequencies which appear in the analysis of a rectifier circuit.

It seems to me that the answer to this question is that these frequencies are not physically present at all, but rather are introduced for convenience in analyzing such a circuit mathematically. It can be shown that the response of a linear network to a periodic, nonsinusoidal wave form can be obtained by adding the responses of the circuit to the various pure sinusoidal components of the actual wave form. These pure sinusoidal components are obtained usually by a Fourier analysis. Since the response of linear networks to sinusoidal waves is well known, this is a convenient procedure, but the frequencies used in the analysis are not physically present in any real sense.

Similar artifices are used in other branches of physics. For instance, we calculate the translatory motion of an extended body by considering all of the mass of the body to be concentrated at its center of mass. Surely no one would claim that the mass of the body is physically located at its center of mass. Similarly, in computing the gravitational attraction between uniform, spherical bodies, all of the mass of each body can be considered to be located at its center as a convenience in making the calculation. Physically, of course, this is not so. Additional examples could easily be given in which the elements of calculation do not correspond literally to the physical situation and are introduced to simplify the analysis.

Franklin and Marshall College,
Lancaster, Pennsylvania

F. W. VAN NAME, JR.

On the Use of Dimensional Analysis in
General Physics

THE subject of dimensional analysis, called the principle of similitude by Raleigh and other English authors, has found great application in theoretical physics and engineering practice, and especially in model experiments. Engineering students are sometimes given goodly doses of it in the senior year, but physics students, beginning or older, by and large, are not usually too well schooled in the matter. General physics textbooks devote little space to it in general. The treatment, if any is given at all, is devoted to showing, in a paragraph or two, how one verifies the homogeneity of an equation, that is, whether the quantities equated have the same physical nature. This is a good thing instructionally, and the advantages of the tool are obvious. I am of the opinion, however, that much more should be done with this device throughout the first physics course, and this note is intended to point up the idea.

The standard and most prevalent type of classroom instruction in general physics finds the professor at the board deriving the equations of the textbook. He may indeed depart from the textbook derivation, and this is good pedagogically, but in my own thinking there is little justification for repeating what is in the book. More emphatically, the teacher should not repeat the book nor reproduce it. Cannot the student read? If a question is raised by an enterprising student on a detail of the proof then, of course, some instruction is needed on the point. But the teacher should bring to the lecture something beyond the textbook. Indeed, the measure of a good textbook lies, in my judgment, in this very quarter. Can the teacher leave the textbook to the student and spend his time teaching what is not in the book? This is an ideal scheme, of course. It appears dipole to the growing block-and-gap method. There is now so much in standard texts that it is often frustrating to make a choice.

But returning to our central theme, the method of establishing a formula by dimensional analysis strikes me as an excellent substitute for the textbook derivations. It can be applied to the simpler expressions first encountered, so that the student gets the feel of the method. As the course advances the more difficult formulas may be established by this method.

The first step, as I point it up, is to "characterize" the problem. By this I mean, what *distinguishing* features does this inquiry possess? What characterizes this problem? This is equivalent to asking: what *variables* might we reasonably expect to enter into this inquiry? Let us take a case to illustrate. A good elementary one is the simple pendulum. Shall we suppose the motion to be *periodic*? This is plausible. We wish, therefore, to find the *law* to which this periodic motion subscribes. Its period T is, then, a function of something. Shall we suppose that the length l of the string is involved? This is reasonable. (But let us disregard its mass.) Will the mass m of the bob

enter? or its weight w perhaps? These may, for all we know. Will the *magnitude of the swing* be involved? This may be, but let us disregard this detail for the moment. It is clear that we have imposed a number of apparently artificial constraints on our argument, but we do not want to make the problem too complicated at the beginning.

We intend then, to find some relation between t , l , m , and w ; that is, $t = f(l, m, w)$. This leads, with appropriate discussion, to writing $t = \text{constant} \times l^a m^b w^c$, and finally to

$$[T] = [L]^a [M]^b [MLT^{-2}]^c.$$

This last is a formal statement of how L , M , and T enter and assures us also of dimensional homogeneity. If one pleases he can take the variables as l , m , and g . The result is the same, obviously. The constant must come from experiment or from a purely mechanical solution.

If, now, we think that the arc of swing s may be involved, we can write $t = \text{constant} \times l^a m^b w^c s^d$, and the student is quite illuminated to discover that s alone is not important; it is rather the ratio of s to l , or the angle, that counts,

Most of the formulas (those which are not pure definition) in general physics lend themselves to this method, although one meets up with some hazard in electricity and magnetism with beginners. (Indeed, one sometimes meets a hazard if he is not a beginner!) My own experience shows that the method is very appealing instructionally.

It possesses an elegance not always found in the usual textbook derivation and the better students find themselves repeatedly drawn to it.

JULIUS SUMNER MILLER

*Dillard University
New Orleans 22, Louisiana*

Focusing with a White Light Source

WHEN a white light source is being used to produce a spectrum for the study of the effects of filters, it is sometimes troublesome to get a sufficiently good focus to insure spectral purity. A very simple and convenient way to check the adjustment of the optical system is to observe the spectrum through didymium glass goggles, such as used in glass blowing. A continuous spectrum observed through these shows a sharp absorption "line" in the yellow when the adjustment is correct. For demonstration purposes, it has been found more effective to let students use the goggles individually than for the instructor to insert them in the light beam.

The same method is, of course, helpful when adjusting a spectroscope.

E. SCOTT BARR

*University of Alabama
University, Alabama*

ANNOUNCEMENTS AND NEWS

Book Reviews

You Will Like Geometry. KARL MENDER. Pp. 34, Figs. 100, 15×22.5 cm. Book Mart, Museum of Science and Industry, Chicago, 1952. Price 10 cents.

This small guide book is keyed to the geometry exhibition arranged by the Illinois Institute of Technology for the Chicago Museum of Science and Industry. One might suppose that it would therefore be dull reading if one could not see the exhibits at the same time. Not so. It is an informative booklet, simple in arrangement, nicely illustrated, and in a format somewhat reminiscent of the Liebers' little volume on *The Education of T. C. Mils*. Its keynote is modernity, and it sharply challenges the general supposition that geometry became a finished product with Euclid in the dim past.

Although the presentation is brief and simple, there are numerous profundities brought to the surface and one is continually led to feel that "what lies over the fence" must be as inviting as what is presented to view. The author asks intriguing questions and displays many of the delightful ingenuities and beauties of his subject. In the field of guide book literature, it deserves a place of eminence. Topics are treated with a brevity and exactness that is admirable. Many old and time-honored geometrical questions are discussed; and in the midst of old friends, one is also certain

to encounter new faces. Just as old friends sometimes present new and unsuspected facets, so will old geometrical friends introduce new and extending aspects, as I found in resuming an earlier interest in the regular solids, and in other parts of Menger's little book.

Anyone for whom geometry is a hobby should certainly have this book. Almost anyone with even a casual interest would find a delightful half-hour in reading it. It not only shows why geometry has fascinated folks for centuries, but it will in turn do its share to maintain and extend that fascination.

RICHARD M. SUTTON
Haverford College

The Exact Sciences in Antiquity. O. NEUGEBAUER. Princeton University Press, 1952. pp. xvi+191, Figs. 44, 16×24.5 cm. Price \$5.00.

Profound changes in the foundations of science during the past half-century have brought home with telling force the persistent need for suspended judgment concerning ultimate principles. If the necessity of similar caution with respect to the history of science has been in danger of going unheeded, the present volume will serve as an antidote for historical complacency. It is, in a sense, an intensely disconcerting work, for the accumulation of new evidence from cuneiform tablets seems but to emphasize how much less one

knows about pre-Hellenistic science than our forefathers thought they knew. Many a cherished dogma, perpetuated through generations of school books, is demolished, often without the mitigating comfort of facile alternatives. The once highly touted role of the Nile valley, buttressed by fables concerning the putative scientific significance of the pyramids, is sharply deflated: "Ancient science was the product of a very few men; and these few happened not to be Egyptians." Gone, too, are the myths concerning an early Sumerian astronomy, the legendary accuracy of Mesopotamian observations, the Babylonian eclipse-predictions based upon the Saros, and the notion that astrology preceded astronomy. The venerated stories of discoveries made by Thales and Pythagoras "must be discarded as totally unhistorical"; and Plato's direct contributions to mathematical knowledge are unceremoniously dismissed as "obviously nil." Neugebauer even goes so far as to deny significance to the ubiquitous division of the history of science into three periods—ancient, medieval, and modern.

In mathematical astronomy ancient methods prevailed until Newton and his contemporaries opened a fundamentally new age by the introduction of dynamics into the discussion of astronomical phenomena. One can perfectly well understand the "Principia" without much knowledge of earlier astronomy, but one cannot read a single chapter in Copernicus or Kepler without a thorough knowledge of Ptolemy's *Almagest*. Up to Newton all astronomy consists in modifications, however ingenious, of Hellenistic astronomy.

One should not conclude that the present book is essentially iconoclastic. Far from it. If recent evidence does make untenable many of the older conventionalities, this is counterbalanced by the disclosure of unanticipated methods and results in Mesopotamian science. One could mention the value of $3\frac{1}{2}$ for π and a calculation to half a dozen places of $\sqrt{2}$; but of more importance is the fact that present knowledge of Babylonian mathematics indicates a level of competence in many respects comparable to that of the early Renaissance. In astronomy one of the most brilliant achievements was the Seleucid recognition of the independence of factors governing the first visibility of the moon and the development of a theory, based upon arithmetic progressions, which permitted the prediction of their combined effect. For the planets also the Babylonians were concerned primarily with horizon phenomena, in contrast to the Ptolemaic interest in geometrical theory; yet it appears that the Babylonian empirical material and arithmetic methods were used in the *Almagest* to devise the geometrical models from which arithmetic consequences were in turn derived.

Apparent throughout the volume is the author's distaste for glittering generalities and his propensity for posing historical problems in all their stark and unsolved simplicity. Nevertheless, the oral form in which the material originally was presented (in a series of lectures at Cornell University in 1949) has encouraged Neugebauer occasionally to throw off inhibitions appropriate to technical publication. Concerning the development of Hellenic geometry,

for example, he makes a provocative conjecture. He believes that early Greek mathematics cannot have been very different from the Heronic Diophantine type, but that an essential turn came in the fifth century B. C. with the discovery of the arithmetic fact of the irrationality of $\sqrt{2}$ and the geometrical corollary concerning the diagonal of a square. Two major steps, both characteristically Greek, seem to have followed: first, the recognition of the need for an axiomatic procedure; and, second, the formulation of arithmetic results in the familiar "geometrical algebra." In this development the "application of areas" would be a natural outgrowth of the Babylonian solution of quadratic equations.

This book is not what its title might imply; it is by no means a compendious account of ancient science, inasmuch as the author all too modestly felt himself "incompetent to deal with medicine or the natural sciences." The emphasis of the six chapters (or lectures) is upon pre-Hellenic mathematics and astronomy, and here Neugebauer goes into considerable technical detail (not always easy to follow) concerning the methods which were used. The concluding chapter, entitled "Origin and Transmission of Hellenistic Science," is devoted to a survey of some historical interrelationships with other ancient civilizations. It is suggested that Babylonian influence in astronomy may have been felt in India through the mediation of pre-Ptolemaic Greek sources.

Besides the text of the lectures, this book contains scholarly apparatus which the serious reader will deeply appreciate—a good index, chronological tables, bibliographies, and (most significant of all) several dozen pages of notes and references. Over-all attractiveness of the volume appropriately complements its authoritative authorship.

CARL B. BOYER
Brooklyn College

Quantum Theory. DAVID BOHM. Pp. 646, Figs. 117, 16×23.5 cm. Prentice-Hall, Inc., New York, 1951. Price \$7.50.

Here is a well-written textbook on quantum mechanics which introduces the student to the subject with primary emphasis on clarity of exposition of the fundamental concepts, and with a corresponding paucity of discussion of applications where many of the startling and convincing numerical successes of the theory are to be found. One spends the first half of the book in the realm of one space dimension, the realm of concepts leading up to the formulation of the wave equation, and useful exercise in the manipulation thereof in simplified one-dimensional applications such as those involving square-well potentials. Into the next third of the book is condensed, mainly by stringent limitation of the number of subjects discussed, the three-dimensional consideration of the stable systems met in physics, including angular momentum and a nonrelativistic treatment of spin. This section also includes the matrix formulation of the theory, but for simplicity only as derived from the wave-mechanical approach. Then the last part of the book contains a relatively discursive treat-

ment of the theory of scattering and ends with some remarks on the theory of measurement.

The emphasis on discussion of the introductory material seems very appropriate for the student at the beginning of his acquaintance with quantum mechanics, and it is interesting to see that the author does not consider a leisurely pace at this point a luxury which one cannot afford. Yet the fact that this rather large textbook of over 600 pages gets no farther than it does with the detailed applications is symptomatic of the pedagogical difficulties in the ever-widening field of physics. Quantum theory now being taken for granted as an indispensable part of the training must be started early in the student's specialized career. It is rather surprising to find in the book indications of how early it is started. The author has a healthy habit of making reference to classical equivalents of the quantum-mechanical problem, and in such a reference to coupled oscillators, no knowledge is assumed or mention made of normal coordinates, useful though this concept would be in the discussion at hand.

At the other end the amount of material left for further courses in quantum mechanics is impressive, and unfortunately perhaps too lengthy to be included in the normal curriculum of the average graduate student. Bohm's book wisely makes no attempt to introduce relativistic quantum mechanics, leaving this to a further course which in many schools might follow Dirac's own terse and elegant presentation, again devoted mainly to simple mechanical systems. But there is also the realm of the more complicated systems, the atoms, molecules, and solids, before coming to the more specialized nuclear field. These are what are given rather short shrift in most general curricula, so that an attempt is frequently made to include a more extensive treatment of them in an early general course in quantum mechanics. It is to be hoped that these subjects may be studied by all students in a separate course, and that the problems of transformation from one extreme representation to another, as typified by the Paschen-Back effect, wherein very impressive comparisons with experiment are possible, may not be neglected. The ideal textbook for such a course has yet to be written, and it would be desirable to have it done with the clarity of presentation and careful selection of material for pedagogical effect characteristic of Bohm's introductory book. Still further training in quantum electrodynamics and field theory may perhaps still be considered the specialized domain of the theoretical student, and, being one step further removed, of course need be considered no concern of the introductory book.

In its discursive treatment of the introductory concepts at the expense of the applications, Bohm's book contrasts strongly with another excellent textbook on the same general subject written by an American author in the last few years, that of Leonard Schiff (McGraw-Hill Book Company, Inc., New York, 1949). Although the latter is only two-thirds as long by page count, it goes a great deal further into the detailed applications in the general fields of radiation, atomic and molecular physics, relativistic quantum mechanics, and quantum electrodynamics. It might thus almost be used as a sequel to Bohm's book, although some discreet elimination of the end of one and the begin-

ning of the other would be required to avoid the sort of repetition which is unfortunate when there is so much of interest that must of necessity be omitted from the general graduate curriculum. In comparing the applications where the two books do overlap, it is to the credit of Bohm's book that, for example, it gives the clearer and more physically understandable explanation of the relation between exchange energy, spin-spin coupling, and the probable distance between the two electrons involved as effected by the symmetry of the wave function. (The subsequent mention of the application to ferromagnetism is at fault in a minor way in implying that the ambiguity in sign enters in the electronic energy, with no mention of the nuclear terms.)

This thoroughly sound book, with its wealth of comparison with classical concepts and consequent bolstering of physical intuition, is to be welcomed as a valuable addition to the family of textbooks that will aid in the training of the next generation of physicists.

DAVID R. INGLIS

Argonne National Laboratory

Theory of Groups and Its Application to Physical Problems.

S. BHAGAVANTAM, AND T. VENKATRAYUDU. Pp. 277 + xiii. 2nd Edition. (In English.) Andhra University, Waltair. Printed at the Bangalore Press, Mysore Road, Bangalore City, 1951. Price Rs. 20.

The theory of groups represents a powerful tool for the solution of physical problems. But unfortunately, too few physicists and chemists have made effective use of group theory methods, probably in large part due to the highly mathematical aspect of books on the subject, many of which have laid little or no stress on the applications to physical problems. The present book is an introductory study of the theory of groups for the scientist who wishes to understand the method and its advantages for physical problems. In this 2nd edition a few alterations and additions have been made.

The first 5 chapters present the elements of group theory, the symmetry properties of 1- and 2-dimensional lattices, and the matrix groups associated with sets of linear transformations. Applications are then made to the wave equation, eigenfunctions of H-like atoms, and to vibrations of a dynamical system. The vibrational Raman and infrared spectra, molecular structure, normal modes, and normal frequencies are considered next. Lattices in 3 dimensions, vibrations of a crystal lattice, Raman scattering in crystals, and rotation groups follow. Finally, applications to atomic spectra, band spectra, crystal symmetry and physical properties, and other applications are given.

In the 8 appendices, group algebra, transformation of matrices, the Kramers-Heisenberg dispersion formula, evaluation of group characters, properties of polynomial functions, the Laplacian operator, reduction of reducible representations, and character tables and irreducible representations of various point groups are treated.

The book gives a connected account of important physical applications of group theory which have hitherto been scattered throughout many journals. It is a book which

can be read with profit by most physicists and chemists, and by many of those in applied science and engineering.

FORREST F. CLEVELAND
Illinois Institute of Technology

Electrical Measurements. FOREST K. HARRIS. Pp. 784.
John Wiley and Sons Inc., New York, 1952. Price \$8.00.

It is customary for authors of books on electrical measurements to write them as texts for undergraduate or graduate courses in the subject and to include in them discussions of experiments suitable for such courses and such other material as may assist a student in understanding the principles of those experiments. Excellent books of this type are available. Dr. Harris' work has no such limited objective. It is rather a complete treatise on the subject of direct-current and low frequency measurements compounded from the author's many years of experience in the field through his work at the National Bureau of Standards and as a university teacher.

The treatment of the nature of precise physical measurements and the methods used to attain them is consistently excellent throughout. Not only the most precise methods of measurement are treated, however, but many are discussed where other criteria obtain. The electrical components of measuring circuits, their alternative constructions, and their errors are exhaustively discussed. The history of the subject is given in each particular section. The history and inter-relations of the different systems of units and the standards of electrical units are presented in lucid detail. Bridges, potentiometers, magnetic measurements, instruments, and components are discussed with a completeness which is made possible by the conciseness of the writing.

The work stops at the point in electrical measurements where vacuum-tube circuits begin to apply. Only a few of the simplest such circuits are mentioned. This may not fairly be taken as a criticism, however, since it would obviously have been impossible to have considered vacuum-tube circuits in addition to the other subject matter and maintained the standards the author has set himself. We can only hope that a book of this stature will be written on the subject of measurements with vacuum-tube circuits.

The typography and illustrations are commendable. The writing style is clear and consistently interesting. It is a "solid" book. Even the footnotes are so informative that it is difficult to see why they were relegated to smaller type. So much good information and advice is included in this book that every serious student or practitioner of experimental physics should read it and, if possible, have it as a ready reference.

W. G. WADEY
Yale University

Basic Kinesiology. FERD JOHN LIPOVETZ. Pp. 105+iii, Figs. 91, 22×28.5 cm. Burgess Publishing Company, Minneapolis, 1952. Price \$3.75.

This "Textbook of Body Mechanics and Analysis of Muscular Movements" appears at a time of increasing

interest in the use of biological material in elementary physics courses, particularly those for premedical students. Many recent textbooks make use of such illustrative material, but further exploration on the part of instructors is complicated by lack of familiarity with the literature. It is so easy to postpone investigation of another department's library!

At first glance this text seems to be the answer to the mechanics part of this problem—the illustrations in sketch and diagram form are clear, and many are accompanied by the explicit numerical material dear to the heart of the problem-loving instructor. Perhaps even more valuable is their suggestion of simple laboratory studies and experiments. The first three chapters are particularly pertinent with their illustrations of anatomy in terms of levers and forces and centers of gravity, and their applications in various sports. The rest of the text is primarily a more detailed analysis of muscular action.

Further investigation unfortunately leads to considerable disillusionment. An awkward style is unfortunately coupled with lack of precision and consistency in the use of the terminology of physics. One meets, for example, "momentum (force any body is driven or impelled)", "we use the expression Power (P) for Force," "the average speed was 3.5 seconds." The correct statement and interesting use of the law of the lever is followed by a "conclusion": "The force required to overcome a resistance varies directly with the length of both the force and weight lever arms." There are elisions of whole sections of the reasoning: for instance in the use of Bernoulli's principle to explain the curving of a ball, there is no statement at all that the velocity of the air with respect to the ball should be considered. It makes physics sound like a series of unrelated and arbitrary laws and formulas.

Much has been said recently about the desirability of acquaintance by the physicist with physiological material. But what about asking a physicist to cooperate in the writing of such books as this? A little expansion in one place, more use of broad underlying principles in another, technical proofreading throughout, seem to be indicated. As it stands, the student who knows a little physics can only be frustrated, the one who knows none be baffled and repelled, by such a presentation. You will find excellent material in it for your own cogitations, but do not be tempted to give it to a student! And if you are new to the field have at hand one of the more complete standard texts;¹ its photographs will be more difficult to interpret in terms of physics but its explanations and definitions will be more complete.

NORA M. MOHLER
Smith College

¹ For example, *Applied Anatomy and Kinesiology*, W. P. Bower, Revised by H. A. Stone (Lea and Febiger, Philadelphia, 1949).

Applied Statistics. A Journal of the Royal Statistical Society. Edited by LEONARD H. C. TIPPETT. Volume 1, No. 1. Pp. 80. 16×25.5 cm. Oliver and Boyd, Ltd., Edinburgh, Scotland. Price 25s (\$4.00) annually.

Applied Statistics was founded "to present, in one way or another but always simply and clearly, the statistical

approach and its value, and to illustrate in original articles modern statistical methods in everyday application."

This first issue is a good example of illustrating how the purposes of this new statistical periodical are to be carried out. The seven articles contain various applications of modern statistical techniques which proved beneficial for solving certain problems. These techniques were successfully applied to industrial situations, management affairs, job evaluation, social classes of towns, association between jury qualification and income status, specification of plastics, sampling from conveyor belts, and periodic sampling.

The authors pointed out the uses of coefficients of variation, analyses of variance, regression equations in several variables, quality control by statistical methods, histograms, frequency distributions, graphical presentations of data, standard deviations, a study of sources of variation, correlation coefficients, indices, factor analyses, systematic sampling, tests of significance between statistics, and also indicated the advantages of knowing some of the underlying principles of statistics.

The section devoted to questions and answers should prove helpful to many, because it offers opportunities for any reader to submit his statistical troubles to expert statisticians who will give correct solutions. This sort of advice should promote a great deal of interest and logical thinking.

Other sections of the journal should keep the reader up to date on the activities of the Society and on some of the latest statistical developments which are presented at meetings at different places. These sections contain abstracts of papers and short summaries of various talks.

The author of the first article is incorrect in the statement "The systematic applications of statistical methods by persons engaged in research work appears to have begun about 1920." Many important research projects were successfully analyzed by Karl Pearson and his school before the third decade of the present century. People from all over the world profited by their research work and statistical developments.

The article on factor analysis is too brief. Anyone not familiar with that phase of statistics will find very little of interest in it.

The appearance of *Applied Statistics* clearly indicates the need for statistical analyses in many fields and also suggests that laymen are clamoring for statistical assistance in their work.

The reviewer believes that this new Journal will meet a pressing demand of laymen in several countries.

WILLIAM D. BATEN

Michigan Agricultural Experiment Station

Intermediate College Mechanics—A Vectorial Treatment.

DAN EDWIN CHRISTIE. Pp. 454+xvi, Figs. 471, 16×23.5 cm. McGraw-Hill Book Company, New York, 1952. Price \$7.00.

Professor Christie's philosophy of teaching mechanics, as expressed in his Preface and Introduction, may be paraphrased in the statement: *Mechanics problems are solved by the proper application of one or more of the relatively small number of physical principles which underlie the subject; the*

student can best understand the principles and their vicissitudes by the practice he gets from doing a large number of problems. With this idea I heartily agree. The student must be taught to make a long-range investment in generality, and eschew the short-term gains that are sometimes gotten from facile substitution of data into specific "formulas." At the several levels of teaching mechanics, the selection of the level of generality of the principles is a subtle and important task; we could hardly expect a Physics-1 student to find how long it would take a baseball to fall from the top of the Washington Monument by using the Principle of Least Action! In classroom practice, the physics teacher should justify the principles intuitively, discuss the concepts contained in them, decompose them into not-necessarily-memorized subprinciples, and illustrate their formal manipulation in actual physical situations.

The first two chapters of this new book discuss vectors through triple product from a purely geometric viewpoint, and (as the subtitle states) a three-dimensional vector notation is used throughout. Other usual mathematical concepts are carefully introduced wherever they are needed. The next nine chapters cover the commonly encountered basic subjects, with six more chapters devoted to more specialized material. The text is formal but lucid, and there is no paucity of diagrams. Problems of varying degrees of complexity, in more than sufficient quantity, are included at several points in each chapter.

The basic concepts (such as time, force, particle) are introduced intuitively. From these are compounded the principles. Important principles are separately stated in italics; but then so are some relatively unimportant corollaries, which attenuate the effect of this device. The principles are illustrated by well-chosen examples. The arguments tend to be logical and mathematically formal rather than intuitive and physical; I would say that the over-all tenor of the book is more mathematical than physical, the author's frequent contentions to the contrary notwithstanding. A simple physical interpretation of many of the *derived* concepts—moment of inertia, radius of gyration, center of pressure, angular momentum, to mention a few—is lacking; these quantities are unearthed and fitted into the structure of mechanics in a formal way. Little space is devoted to *mechanisms*; for example, the experimental genesis of the concept of coefficient of friction is not mentioned. The mathematical foundation is built far stronger than the superstructure that it must immediately support; the vectorial analysis is three-dimensional, but virtually all of the problems dealing with physical situations are two-dimensional. This effect is commonly observed in textbooks; it is not necessarily a drawback. In short, the *physical principle + practice* concept is implied rather than stressed throughout the text.

The teachability of the book is improved by many novel features. Vectors are always printed in boldfaced capitals and scalars in lower case type; thus, the teacher at the blackboard and the student with his notebook are relieved of the necessity of putting special identifying marks on vector quantities. Force, length, and time are taken as the basic ingredients of mechanics; the elusive (even at this level!) quantity mass is defined in terms of them, and a set of

problems on units is included. There is a section on *vectorial* dimensional analysis, using f , l , t , and a unit vector. Couples are brought up to their rightful position in the study of mechanics. The author frequently repeats (and practices) the deathless injunction *isolate the system*; he includes a set of problems just to give the student practice in doing so. The only serious misprint is in the blurb on the dust-cover. Coriolis forces are covered thoroughly, though nonmechanistically; the dubious proposition that this effect causes spiraling of water in a sink drain is stated.

The presentation is logical and even, with the vectorial

treatment making for smooth sailing through many usually messy topics. The derivation of the conic section paths of particles in an inverse-square field falls out quite easily, and the chapter on Deformable Bodies in Equilibrium is especially good. Professor Christie does not hesitate to stop to fill in blind spots in the student's early training. His book is sure to attain a competitive position among the half-dozen most popular texts on the subject of intermediate mechanics.

IRVING L. KOFSKY
Smith College

RECENT MEETINGS

American Society for Engineering Education Physics Division

The Physics Division of the American Society for Engineering Education met with the parent society at Dartmouth College, June 23-27, 1952. The program consisted of a joint conference with the Division of Educational Methods, a luncheon with the Division of Mechanical Engineering, participation in a conference on Atomic Energy in Engineering Education, and an afternoon's conference on modern physics in the engineering curriculum. It is hoped that the good attendance at these meetings marks an increasing recognition of the desirability of improved liaison between those who teach physics and those who teach its applications in the various fields of engineering.

The following officers were elected: *Division Chairman*, G. P. BREWINGTON, *Lawrence Institute of Technology*; *Division Vice-Chairman*, RAYMOND J. SEEGER, *National Science Foundation*; *Division Secretary*, GEORGE H. BURNHAM, *Norwich University*; *Council Representative*, ELMER HUTCHISSON, *Case Institute of Technology*; *Board of Directors*, C. E. BENNETT, *University of Maine*, HOMER L. DODGE, *Norwich University*, J. G. POTTER, *Agricultural and Mechanical College of Texas*, and MARSH W. WHITE, *The Pennsylvania State College*.

The following papers were presented:

The nature of a continuing and effective society program for improving teaching. Panel discussion. R. L. SWEIGERT, *Georgia Institute of Technology*, Chairman; N. A. CHRISTIANSEN, *Cornell University*; N. W. DOUGHERTY, *University of Tennessee*; ERNST WEBER, *Brooklyn Polytechnic Institute*; B. R. TEARE, *Carnegie Institute of Technology*; H. P. HAMMOND, *Pennsylvania State College*.

The year of evaluation. Open discussion; E. E. AMBROSIUS, *Pennsylvania State College*, presiding.

Report of April meeting of Steering Committee of ASEE Committee on Atomic Energy Education. ALBERTO THOMPSON, *Atomic Energy Commission*.

Engineering courses bearing applications of nuclear energy. ROBERT C. ERNST, *University of Louisville*.

The education of nuclear engineers as seen by an AEC contractor. D. W. MCLENNAN, *General Electric Company*.

The AEC looks at the problem of supplying nuclear reactors for engineering colleges. T. KEITH GLENNAN, *U. S. Atomic Energy Commission*.

The modern physics course at MIT. FRANCIS W. SEARS, *Massachusetts Institute of Technology*.

What should the engineer be taught about the theory of the solid state? J. B. FISK, *Bell Telephone Laboratories*.

The training of undergraduates in nucleonics. W. R. KANNE, *General Electric Company*.

GEORGE H. BURNHAM, *Secretary*

The gulf between percepts and physics is not a gulf as regards intrinsic quality, for we know nothing of the intrinsic quality of the physical world, and therefore do not know whether it is, or is not, very different from that of percepts. The gulf is as to what we know about the two realms. We know the quality of percepts, but we do not know their laws so well as we could wish. We know the laws of the physical world, in so far as these are mathematical, pretty well, but we know nothing else about it. If there is any intellectual difficulty in supposing that the physical world is intrinsically quite unlike that of percepts, this is a reason for supposing that there is not this complete unlikeness. And there is a certain ground for such a view, in the fact that percepts are part of the physical world, and are the only part that we can know without the help of rather elaborate and difficult inferences.—BERTRAND RUSSELL, *The Analysis of Matter* (1927).